

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-304360

(43)Date of publication of application : 13.11.1998

(51)Int.Cl.

H04N 7/24
G10L 7/04
// H03M 7/30

(21)Application number : 09-281498

(71)Applicant : MATSUSHITA ELECTRIC IND CO LTD

(22)Date of filing : 15.10.1997

(72)Inventor : MATSUMOTO TAKAO

YONEDA AKI
HORIUCHI KOICHI
TATSUMI HIDENORI
KAWAHARA EIJI
ARASE YOSHITAKA

(30)Priority

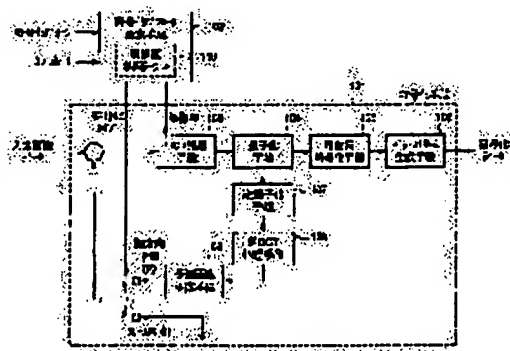
Priority number : 08272746	Priority date : 15.10.1996	Priority country : JP
08284353	25.10.1996	JP
08314563	26.11.1996	JP
09 42051	26.02.1997	JP

(54) METHOD FOR ENCODING VIDEO/VOICE, DEVICE THEREFOR AND MEDIUM FOR RECORDING CODED PROGRAM

(57)Abstract:

PROBLEM TO BE SOLVED: To properly set an encoding parameter corresponding to the basic capability of a computer by determining the encoding parameter based on at least one of resolution, frame rate, and encoding parameter influencing processing performance.

SOLUTION: This device is provided with an encoding means 101 and an encoding parameter decision means 102 including a resolution reference table 110. The encoding parameter decision means 102 determines resolution by using the included resolution reference table 110 from a designated frame rate and an encoding pattern, and outputs the encoding parameter including a parameter indicating the determined resolution to the encoding means 101. The encoding means 101 operates an encoding processing according to the encoding parameter so that encoding with higher resolution can be attained while satisfying a requested condition. The encoding pattern can be determined corresponding to the designated resolution by using the similar reference table.



*** NOTICES ***

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] As opposed to original video information by which an image was digitized in a video encoding method which codes an image and which consists of two or more still picture information, An image coding step coded according to an encoding parameter which mentions one or more of the above-mentioned still picture information later, A frame rate demanded when reproducing resolution which original video information has, and coding data obtained by coding, Processing performance which shows throughput of coding equipment which performs the above-mentioned image coding step, Or a video encoding method performing an encoding parameter determination step which determines the one or more above-mentioned encoding parameters based on one which influences a throughput of coding processing in the above-mentioned image coding step, or any 1 or more of two or more encoding parameters.

[Claim 2] A video encoding method judging throughput of coding equipment which performs the above-mentioned image coding step of the video encoding method concerned in the video encoding method according to claim 1, and performing further a throughput judgment step which outputs a decision result.

[Claim 3] In the video encoding method according to claim 1 or 2, the above-mentioned encoding parameter, A video encoding method being a thing containing one or more of detection ranges at the time of detecting a coding type in which resolution in coding processing performed to the above-mentioned field video information, formation of a frame inner code, or prediction coding is shown, or a motion vector used for the above-mentioned prediction coding.

[Claim 4] A video encoding method being what makes the above-mentioned judgment by the above-mentioned throughput judgment step in the video encoding method according to claim 2 based on a kind of control device which the video encoding method concerned has.

[Claim 5] A video encoding method characterized by being what makes the above-mentioned judgment based on the time required of coding processing in the above-mentioned coding step by the above-mentioned throughput judgment step in the video encoding method according to claim 2.

[Claim 6] In the video encoding method according to claim 2, in the above-mentioned throughput judgment step. Accumulate the above-mentioned original video information inputted temporarily, and in this accumulation, While saving a series of still picture information which constitutes the above-mentioned field video information one by one, An image buffering step which discards still picture information to which it was read in the above-mentioned coding step, and the above-mentioned coding processing was carried out one by one, Preservation of a series of above-mentioned still picture information in the above-mentioned image buffering step, A frame rate control step controlled to carry out in a fixed frame rate determined based on a frame rate which was given as for the account of the upper is performed, A video encoding method being what makes the above-mentioned judgment based on an accumulated dose of the above-mentioned field video information accumulated in the above-mentioned image buffering step temporarily.

[Claim 7] A sound encoding method which codes with a zone part tally item-ized method to a sound, comprising: Set frequency f_s which is a numerical value used for coding processing.

A memory step which memorizes the transformation constant n .

A voice input step which inputs a sound which is an object of coding.

A voice inputting sampling step which creates sampling sound voice data using a sampling frequency determined based on the set frequency f_s which memorized [above-mentioned], It is considered as the m number of sampling sound voice data obtained when the above-mentioned set frequency f_s is used as a sampling frequency, a number defined based on the above-mentioned transformation constant n is made into m' , and it is sampling

sound voice data of m' individual.

[Claim 8] In the sound encoding method according to claim 7, in the above-mentioned voice inputting sampling step. By audio sampling processing inputted [above-mentioned] by making into a sampling frequency set frequency f_s which memorized [above-mentioned]. Create m sampling sound voice data and in the above-mentioned voice data converting step. From the m above-mentioned sampling sound voice data, extract sampling sound voice data every individual $(n-1)$, and between two adjoining sampling sound voice data which extracted [above-mentioned], $(n-1)$ A sound encoding method being what inserts voice data of an individual and changed into m conversion voice data.

[Claim 9] A sound encoding method being what creates conversion voice data in which sampling sound voice data which extracted [above-mentioned] continues n pieces at a time by the above-mentioned voice data converting step in the sound encoding method according to claim 8, respectively.

[Claim 10] In the sound encoding method according to claim 7, in the above-mentioned voice inputting sampling step. By audio sampling processing inputted [above-mentioned] by making into a sampling frequency frequency f_s/n obtained from the set frequency f_s which memorized [above-mentioned], and the transformation constant n . Create sampling sound voice data of a m/n individual, and in the above-mentioned voice data converting step. A sound encoding method being what inserts voice data of an individual $(n-1)$ between two adjoining sampling sound voice data, and changed into m conversion voice data based on the above-mentioned sampling sound voice data.

[Claim 11] A sound encoding method being that in which sampling sound voice data of the above-mentioned m/n individual creates conversion voice data which continues n pieces at a time, respectively by the above-mentioned voice data converting step in the sound encoding method according to claim 10.

[Claim 12] A voice buffering step which holds the above-mentioned sampling sound voice data temporarily to an input buffer in the sound encoding method according to any one of claims 7 to 11, Investigate data volume of the above-mentioned input buffer, and it compares with a value which set this up beforehand, Based on a result of the above-mentioned comparison, perform an input buffer surveillance step which changes a value of the above-mentioned transformation constant n memorized by the above-mentioned register, and in the above-mentioned voice inputting sampling step. A sound encoding method which writes the above-mentioned sampling sound voice data in the above-mentioned input buffer, reads sampling sound voice data from the above-mentioned input buffer in the above-mentioned voice data converting step, and is characterized by being what carries out the above-mentioned conversion of this.

[Claim 13] In the sound encoding method according to any one of claims 7 to 11, the amount of coding data per unit time outputted in the above-mentioned coding step is investigated, A sound encoding method being what performs a coding data surveillance step which changes a value of the above-mentioned transformation constant n memorized by the above-mentioned register based on a result of the above-mentioned comparison as compared with a value which set this up beforehand.

[Claim 14] In a sound encoding method which codes to a sound using a zone part tally item-ized method, Sampling processing of a control parameter memory step which memorizes a control parameter used for the above-mentioned coding, and the voice inputting is carried out, A sampling step which outputs sampling data, and a zone division step which performs zone division to sampling data obtained at the above-mentioned sampling step, and outputs zone signal data, An encoded bit quota step which assigns an encoded bit to zone signal data obtained by the above-mentioned zone division step, A quantization step which quantizes the above-mentioned zone signal data and outputs a quantized value according to assignment of the above-mentioned encoded bit, A coding step which outputs coding data based on a quantized value acquired by the above-mentioned quantization step, A sound encoding method performing the above-mentioned zone division step, the above-mentioned encoded bit quota step, the above-mentioned quantization step, and a coding processing control step that controls data processing in the above-mentioned coding step based on a control parameter which memorized [above-mentioned].

[Claim 15] In the sound encoding method according to claim 14, in the above-mentioned control parameter memory step. As the above-mentioned control parameter, memorize the unit time period judging constant k to an unit time period judging constant register, and the above-mentioned coding processing control step, Set the target number of sampling data to p by one zone split application in the above-mentioned zone division step, and time to be equivalent to p sampling data is made into an unit time period, That judgment which a corresponding unit time period is a coding subject period, or is a period outside a coding subject every p of the above-

mentioned sampling data outputted, Only when it carries out based on an unit time period judging constant which memorized [above-mentioned] and the above-mentioned unit time period is judged to be the above-mentioned coding subject period, When it controls so that sampling data of this unit time period are outputted to the above-mentioned zone division step, and the above-mentioned unit time period is judged to be the above-mentioned period outside a coding subject, A sound encoding method being a judgment control step controlled to output fixed coding data memorized beforehand as coding data in the above-mentioned coding step.

[Claim 16]In the sound encoding method according to claim 15, in the above-mentioned judgment control step. A sound encoding method being what judged as the above-mentioned unit time period t_i being the above-mentioned coding subject period when $i=nxk+1$ is materialized from the unit time period judging constant k which memorized [above-mentioned] by setting the i -th unit time period to t_i , and the arbitrary integers n .

[Claim 17]In the sound encoding method according to claim 14, in the above-mentioned control parameter memory step. As the above-mentioned control parameter, memorize the data-processing judging constant q to a data-processing judging constant register, and the above-mentioned coding processing control step, A sound encoding method being a data-processing step which it is included by the above-mentioned zone division step, and is controlled based on the data-processing judging constant q which memorized [above-mentioned] to close data processing in the above-mentioned zone division step on the way.

[Claim 18]A sound encoding method being what controls data processing of a basic low pass filter in the above-mentioned zone division step by the above-mentioned data-processing step in the sound encoding method according to claim 17 to close on the way about a part for a both-ends step of this filter.

[Claim 19]In the sound encoding method according to claim 14, in the above-mentioned control parameter memory step. As the above-mentioned control parameter, memorize the zone selected constant r to a zone selected-constant register, and the above-mentioned coding processing control step, It receives seeing, although it chose based on the zone selected constant r which memorized [above-mentioned] among zone signal data which the above-mentioned zone division step outputs, A sound encoding method being a zone infanticide step controlled to perform processing in the above-mentioned encoded bit quota step and the above-mentioned quantization step.

[Claim 20]A sound encoding method being what chooses zone signal data from a zone signal data output of M pieces obtained by the above-mentioned zone division step at the above-mentioned zone infanticide step in the sound encoding method according to claim 19 every [which is the zone selected constant which memorized / above-mentioned] r pieces.

[Claim 21]A sound encoding method which acquires a situation of data processing in voice coding, and is characterized by being what performs a processing situation surveillance step which changes a value of the above-mentioned control parameter which memorized [above-mentioned] according to a this acquired situation in the sound encoding method according to any one of claims 14 to 20.

[Claim 22]In the sound encoding method according to claim 21, in the above-mentioned processing situation surveillance step. A voice buffering step which accumulates sampling data in an input buffer temporarily, A sound encoding method being what performs an input surveillance step which makes the above-mentioned control parameter change based on a result of the above-mentioned comparison as compared with a value which set up beforehand quantity of data held at the above-mentioned input buffer.

[Claim 23]In the sound encoding method according to claim 21, the above-mentioned processing situation surveillance step, A sound encoding method characterized by being a coding surveillance step which changes a value of the above-mentioned control parameter based on a result of the above-mentioned comparison as compared with a value which set up beforehand quantity of the above-mentioned coding data outputted per unit time in the above-mentioned coding step.

[Claim 24]In a sound encoding method which codes to fundamental tone voice information by which a sound was digitized using a zone part tally item-ized method, A sampling step which carries out sampling processing of the voice inputting, and outputs sampling data, A zone division step which performs zone division to sampling data obtained at the above-mentioned sampling step, and outputs zone signal data, An encoded bit quota step which assigns an encoded bit to zone signal data obtained by the above-mentioned zone division step, A bit assignment control step which controls assignment in the above-mentioned encoded bit quota step by a mental acoustic-sense analysis substitution control method, A sound encoding method performing a quantization step which quantizes the above-mentioned zone signal data and outputs a quantized value, and a coding step which outputs coding data based on a quantized value acquired by the above-mentioned quantization step according to assignment of the above-mentioned encoded bit.

[Claim 25]In the sound encoding method according to claim 24, the above-mentioned bit assignment control step, A sound encoding method, wherein it controls that it is a bit assignment step one by one according to the order of bit assignment beforehand defined by a mental acoustic-sense analysis substitution control method to zone signal data obtained by the above-mentioned zone division step to perform encoded bit assignment.

[Claim 26]In the sound encoding method according to claim 24, the above-mentioned bit assignment control step, Weighting to each zone beforehand defined by a mental acoustic-sense analysis substitution control method to zone signal data obtained by the above-mentioned zone division step, A sound encoding method being a zone output adaptation bit assignment step controlled to perform encoded bit assignment based on an output level which each zone signal data has.

[Claim 27]In the sound encoding method according to claim 24, the above-mentioned bit assignment control step, Weighting to each zone beforehand defined by a mental acoustic-sense analysis substitution control method to zone signal data obtained by the above-mentioned zone division step, A sound encoding method being an advanced zone output adaptation bit assignment step controlled to perform encoded bit assignment based on weighting to the number of bit assignment for every zone, and an output level which each zone signal data has.

[Claim 28]In the sound encoding method according to claim 24, the above-mentioned bit assignment control step, Comparison with an audibility threshold value is performed for every zone signal data to zone signal data obtained by the above-mentioned zone division step, A sound encoding method being an audibility threshold comparison step controlled not to perform bit assignment to zone signal data judged by the above-mentioned comparison to be less than an audibility threshold, but to increase bit assignment to other zones to it.

[Claim 29]In a video voice encoding method which performs a part or all of a processing process that is included in two above-mentioned coding processings using a common computer resource in coding an image and a sound, When video audio information which comprises original video information which consists of two or more still picture information showing a still picture for every unit time, and fundamental tone voice information showing a sound is inputted, A voice buffering step which accumulates the above-mentioned fundamental tone voice information temporarily, A voice coding step which reads fundamental tone voice information accumulated in the above-mentioned voice buffering step, carries out coding processing of this read above-mentioned fundamental tone voice information, and outputs coding audio information, Throughput about the video voice coding processing concerned is judged using coding load standard information showing a load grade of image coding, A coding load evaluation step which controls coding to original video information in an image coding step mentioned later based on a result of the judgment, A video voice encoding method performing an image coding step which carries out coding processing of the still picture information which constitutes the inputted above-mentioned field video information according to control in the above-mentioned coding load evaluation step, and outputs encoded video information.

[Claim 30]In the video voice encoding method according to claim 29, the above-mentioned coding load evaluation step, A total amount of fundamental tone voice information accumulated in the above-mentioned voice buffering step when still picture information which constitutes the above-mentioned field video information was inputted, Based on the above-mentioned coding load standard information, search for coding load evaluation information, and the above-mentioned coding load evaluation information is compared with a load limit set up beforehand, A video voice encoding method being what cancels the above-mentioned still picture information when still picture information is outputted when the above-mentioned coding load evaluation information has not reached the above-mentioned load limit, and the above-mentioned coding load evaluation information reaches the above-mentioned load limit.

[Claim 31]When video resolution information which inputs and mentions analog video information later in the video voice encoding method according to claim 29 is outputted, Consist the above-mentioned analog video information of two or more discrete digital pixel information, and it changes into original video information which comprises two or more still picture information with resolution according to the above-mentioned video resolution information, Perform an image capture step to output so that it may be processed in the above-mentioned image coding step, and in the above-mentioned coding load evaluation step. A total amount of fundamental tone voice information accumulated in the above-mentioned voice buffering step, Coding load evaluation information is searched for based on coding load standard information showing a load grade of image coding, Based on the above-mentioned coding load evaluation information, search for video resolution information showing resolution of an image used for image coding, output the above-mentioned video resolution information, and in the above-mentioned image coding step. A video voice encoding method being what performs coding processing to the above-mentioned still picture information according to the above-mentioned video resolution

information, and outputs encoded video information when the above-mentioned video resolution information is outputted.

[Claim 32] In the video voice encoding method according to claim 29, in the above-mentioned coding load evaluation step. Output coding load evaluation information so that it may be processed in the above-mentioned image coding step, and in the above-mentioned image coding step. A video voice encoding method being what only a throughput calculated to the above-mentioned still picture information using coding load evaluation information outputted [above-mentioned] performs coding processing, and outputs it as encoded video information.

[Claim 33] In the video voice encoding method according to any one of claims 29 to 32, in the above-mentioned voice coding step. Fundamental tone voice information accumulated in the above-mentioned voice buffering step is read, Calculate a total amount of this read above-mentioned fundamental tone voice information, output as an amount of processed speech information, carry out coding processing of the above-mentioned fundamental tone voice information after that, output as coding audio information, and in the above-mentioned coding load evaluation step Lapsed time, Fundamental tone voice input quantity is calculated based on input quantity per time of the above-mentioned fundamental tone voice information, A video voice encoding method being what calculates the amount of prediction-voices buffers which is a difference of this fundamental tone voice input quantity and the above-mentioned amount of processed speech information, and searches for the above-mentioned coding load evaluation information using the above-mentioned amount of prediction-voices buffers.

[Claim 34] In the video voice encoding method according to any one of claims 29 to 32, in the above-mentioned coding load evaluation step. When the above-mentioned still picture information is inputted, fundamental tone voice input quantity is calculated based on lapsed time and input quantity per time of the above-mentioned fundamental tone voice information, And the amount of processed speech information is calculated based on a total amount of coding audio information outputted in the above-mentioned voice coding step, A video voice encoding method the account of the upper being [fundamental tone voice input quantity which the account of upper calculated, and] what searches for the above-mentioned coding load evaluation information using the above-mentioned amount of prediction-voices buffers after calculating the amount of prediction-voices buffers which is a difference with the calculated amount of processed speech information.

[Claim 35] A video voice encoding method supervising change as a result of the above-mentioned judgment in the above-mentioned coding load evaluation step, and setting up the above-mentioned coding load standard information corresponding to the above-mentioned change in the video voice encoding method according to any one of claims 29 to 34.

[Claim 36] Image coding equipment which codes an image, comprising:

An image encoding means coded according to an encoding parameter which mentions one or more of the above-mentioned still picture information later to original video information by which an image was digitized, and which consists of two or more still picture information.

Make one or more resolution into an encoding parameter of 1, and Formation of a frame inner code, One or more coding types among coding types including each type of forward direction prediction coding, opposite direction prediction coding, and bidirectional prediction coding as other encoding parameters, An encoding parameter deciding means which determines an encoding parameter which is what determines a throughput of the above-mentioned encoding means based on a frame rate which was able to be given.

[Claim 37] A voice to digital converter which codes with a zone part tally item-ized method to a sound, comprising:

Set frequency f_s which is a numerical value used for coding processing.

A register which memorizes the transformation constant n .

A voice input means which inputs a sound which is an object of coding.

A voice inputting sampling means which creates sampling sound voice data using a sampling frequency determined based on the set frequency f_s which memorized [above-mentioned].

Consider it as the m number of sampling sound voice data obtained when the above-mentioned set frequency f_s is used as a sampling frequency, and a number defined based on the above-mentioned transformation constant n is made into m' , A voice data conversion method which outputs conversion voice data containing sampling sound voice data of m' individual which consists of m voice data.

A zone division means to carry out zone division of the above-mentioned conversion voice data, and to acquire M zone signals.

Frequency $f_s / 2n$ obtained from the set frequency f_s which memorized [above-mentioned], and the transformation constant n are made into limit frequency, An encoded bit allocation means which assigns an encoded bit only to a zone signal below limit frequency among the above-mentioned zone signals, A quantization means which quantizes based on an encoded bit which the account of the upper assigned, an encoding means which outputs data which quantized [above-mentioned] as coding data, and a coding data recording device which records the above-mentioned coding data outputted.

[Claim 38] A voice to digital converter which codes to a sound using a zone part tally item-ized method, comprising:

A control parameter memory measure which memorizes a control parameter used for the above-mentioned coding.

A sampling means which carries out sampling processing of the voice inputting, and outputs sampling data.

A zone division means to perform zone division to sampling data obtained by the above-mentioned sampling means, and to output zone signal data.

An encoded bit allocation means which assigns an encoded bit to zone signal data obtained by the above-mentioned zone division means.

A quantization means which quantizes the above-mentioned zone signal data and outputs a quantized value according to assignment of the above-mentioned encoded bit.

An encoding means which outputs coding data based on a quantized value acquired by the above-mentioned quantization means.

A coding processing controlling means which controls data processing in the above-mentioned zone division means, the above-mentioned encoded bit allocation means, the above-mentioned quantization means, and the above-mentioned encoding means based on a control parameter which memorized [above-mentioned].

[Claim 39] A voice to digital converter which codes to a sound using a zone part tally item-ized method, comprising:

A sampling means which carries out sampling processing of the voice inputting, and outputs sampling data.

A zone division means to perform zone division to sampling data obtained by the above-mentioned sampling means, and to output zone signal data.

An encoded bit allocation means which assigns an encoded bit to zone signal data obtained by the above-mentioned zone division means.

A bit assignment control means which controls assignment in the above-mentioned encoded bit allocation means by a mental acoustic-sense analysis substitution control method.

A quantization means which quantizes the above-mentioned zone signal data and outputs a quantized value according to assignment of the above-mentioned encoded bit.

An encoding means which outputs coding data based on a quantized value acquired by the above-mentioned quantization means.

[Claim 40] Video voice coding equipment which performs a part or all of a processing process that is included in two above-mentioned coding processings using a common computer resource in coding an image and a sound characterized by comprising the following.

Original video information which consists of two or more still picture information showing a still picture for every unit time.

A voice buffering means to accumulate the above-mentioned fundamental tone voice information temporarily when video audio information which comprises fundamental tone voice information showing a sound is inputted.

A voice coding means to read fundamental tone voice information accumulated in the above-mentioned voice buffering means, to carry out coding processing of this read above-mentioned fundamental tone voice information, and to output coding audio information.

Coding load evaluation methods which control an output of the above-mentioned field video information over an image encoding means which judges throughput of the video voice coding equipment concerned using coding load standard information showing a load grade of image coding, and is later mentioned based on a result of the judgment.

An image encoding means which carries out coding processing of the above-mentioned still picture information, and outputs encoded video information when still picture information which constitutes the above-mentioned

field video information is inputted according to control of the above-mentioned coding load evaluation methods.

[Claim 41] A recording medium which recorded an image encoding program which carries out coding processing of the image, comprising:

An image coding step coded according to an encoding parameter which mentions one or more of the above-mentioned still picture information later to original video information by which an image was digitized, and which consists of two or more still picture information.

One or more resolution is made into an encoding parameter of 1, and it is each type of formation of a frame inner code, forward direction prediction coding, opposite direction prediction coding, and bidirectional prediction coding.

[Claim 42] A recording medium which recorded a voice coding program which codes with a zone part tally itemized method to a sound, comprising:

Set frequency f_s which is a numerical value used for coding processing.

A memory step which memorizes the transformation constant n .

A voice input step which inputs a sound which is an object of coding.

A voice inputting sampling step which creates sampling sound voice data using a sampling frequency determined based on the set frequency f_s which memorized [above-mentioned], It is considered as the m number of sampling sound voice data obtained when the above-mentioned set frequency f_s is used as a sampling frequency, a number defined based on the above-mentioned transformation constant n which is $m \geq m'$ is made into m' , and it is sampling sound voice data of m' individual.

[Claim 43] In a recording medium which recorded a voice coding program which codes using a zone part tally itemized method to a sound, Sampling processing of a control parameter memory step which memorizes a control parameter used for the above-mentioned coding, and the voice inputting is carried out, A sampling step which outputs sampling data, and a zone division step which performs zone division to sampling data obtained at the above-mentioned sampling step, and outputs zone signal data, An encoded bit quota step which assigns an encoded bit to zone signal data obtained by the above-mentioned zone division step, A quantization step which quantizes the above-mentioned zone signal data and outputs a quantized value according to assignment of the above-mentioned encoded bit, A coding step which outputs coding data based on a quantized value acquired by the above-mentioned quantization step, Based on a control parameter which memorized [above-mentioned], the above-mentioned zone division step, the above-mentioned encoded bit quota step, A voice coding program recording medium recording an encoding program which performs the above-mentioned quantization step and a coding processing control step which controls data processing in the above-mentioned coding step.

[Claim 44] In a recording medium which recorded a voice coding program which codes using a zone part tally itemized method to a sound, A sampling step which carries out sampling processing of the voice inputting, and outputs sampling data, A zone division step which performs zone division to sampling data obtained at the above-mentioned sampling step, and outputs zone signal data, An encoded bit quota step which assigns an encoded bit to zone signal data obtained by the above-mentioned zone division step, A bit assignment control step which controls assignment in the above-mentioned encoded bit quota step by a mental acoustic-sense analysis substitution control method, A quantization step which quantizes the above-mentioned zone signal data and outputs a quantized value according to assignment of the above-mentioned encoded bit, A voice coding program recording medium recording an encoding program which performs a coding step which outputs coding data based on a quantized value acquired by the above-mentioned quantization step.

[Claim 45] In a recording medium which recorded a video voice encoding program which performs a part or all of a processing process that is included in two above-mentioned coding processings using a common computer resource in coding an image and a sound, When video audio information which comprises original video information which consists of two or more still picture information showing a still picture for every unit time, and fundamental tone voice information showing a sound is inputted, A voice buffering step which accumulates the above-mentioned fundamental tone voice information temporarily, A voice coding step which reads fundamental tone voice information accumulated in the above-mentioned voice buffering step, carries out coding processing of this read above-mentioned fundamental tone voice information, and outputs coding audio information, Throughput about the video voice coding processing concerned is judged using coding load standard information showing a load grade of image coding, A coding load evaluation step which controls coding to original video information in an image coding step mentioned later based on a result of the judgment, A video voice encoding

program recording medium recording an encoding program which performs an image coding step which carries out coding processing of the still picture information which constitutes the inputted above-mentioned field video information according to control in the above-mentioned coding load evaluation step, and outputs encoded video information.

[Translation done.]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-304360

(43) 公開日 平成10年(1998)11月13日

(51) Int.Cl.⁶

識別記号

F I

H 0 4 N 7/24

H 0 4 N 7/13

Z

G 1 0 L 7/04

G 1 0 L 7/04

G

// H 0 3 M 7/30

H 0 3 M 7/30

Z

審査請求 未請求 請求項の数45 O L (全114頁)

(21) 出願番号 特願平9-281498

(22) 出願日 平成9年(1997)10月15日

(31) 優先権主張番号 特願平8-272746

(32) 優先日 平8(1996)10月15日

(33) 優先権主張国 日本 (J P)

(31) 優先権主張番号 特願平8-284353

(32) 優先日 平8(1996)10月25日

(33) 優先権主張国 日本 (J P)

(31) 優先権主張番号 特願平8-314563

(32) 優先日 平8(1996)11月26日

(33) 優先権主張国 日本 (J P)

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 松本 孝夫

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72) 発明者 米田 亜旗

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72) 発明者 堀内 浩一

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(74) 代理人 弁理士 早瀬 憲一

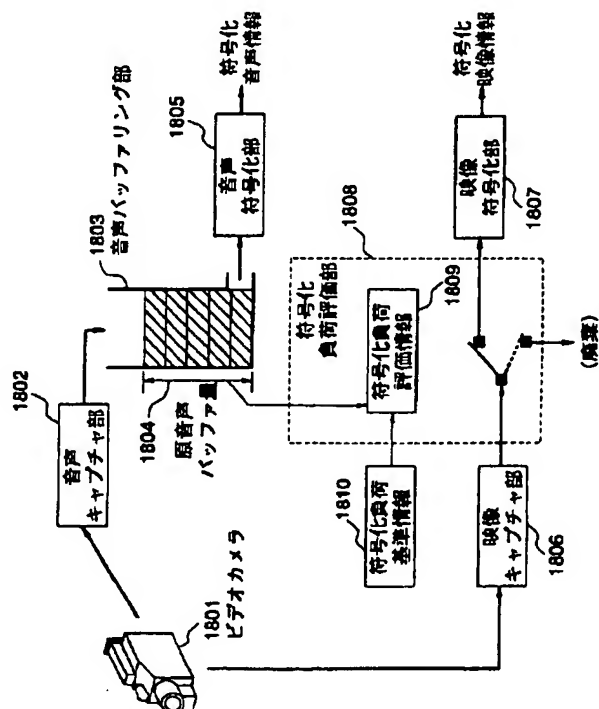
最終頁に続く

(54) 【発明の名称】 映像・音声符号化方法、符号化装置、及び符号化プログラム記録媒体

(57) 【要約】

【課題】 ソフトウェアによる映像、音声、映像音声符号化処理において、ハードウェア的環境と、ソフトウェア的環境に対応して、符号化条件を決定し、装置資源を活用して、良好な符号化結果の得られる符号化方法を提供する。

【解決手段】 ビデオカメラ1801からとりこんだ映像音声を符号化するにあたり、デジタル化した音声データを音声バッファリング部1803に一時蓄積し、当該蓄積量1804に応じて、デジタル化した映像データを符号化処理するか、処理せずに廃棄するかを決定する。



【特許請求の範囲】

【請求項1】 映像を符号化する映像符号化方法において、

映像がデジタル化された、複数の静止画像情報からなる原映像情報に対して、上記静止画像情報の1つまたは複数、後述する符号化パラメータに従って符号化する映像符号化ステップと、

原映像情報の有する解像度、符号化によって得られる符号化データを再生する際に要求されるフレームレート、上記映像符号化ステップを実行する符号化装置の処理能力を示す処理性能、または上記映像符号化ステップにおける符号化処理の処理量に影響する1つ、もしくは複数の符号化パラメータのうちいずれか1つ以上に基づいて、1つ以上の上記符号化パラメータを決定する符号化パラメータ決定ステップとを実行することを特徴とする映像符号化方法。

【請求項2】 請求項1に記載の映像符号化方法において、

当該映像符号化方法の、上記映像符号化ステップを実行する符号化装置の処理能力を判断して、判断結果を出力する処理能力判断ステップをさらに実行することを特徴とする映像符号化方法。

【請求項3】 請求項1または2に記載の映像符号化方法において、

上記符号化パラメータは、上記原映像情報に対して行う符号化処理における解像度、フレーム内符号化、もしくは予測符号化を示す符号化タイプ、または上記予測符号化に用いる動きベクトルを検出する際の検出範囲のうち1つ以上を含むものであることを特徴とする映像符号化方法。

【請求項4】 請求項2に記載の映像符号化方法において、

上記処理能力判断ステップでは、当該映像符号化方法の有する制御装置の種類に基づいて上記判断を行うものであることを特徴とする映像符号化方法。

【請求項5】 請求項2に記載の映像符号化方法において、

上記処理能力判断ステップでは、上記符号化ステップにおける符号化処理の所要時間に基づいて上記判断を行うものであることを特徴とする映像符号化方法。

【請求項6】 請求項2に記載の映像符号化方法において、

上記処理能力判断ステップでは、上記入力される原映像情報を一時蓄積し、該蓄積にあたっては、上記原映像情報を構成する一連の静止画像情報を順次保存していくとともに、上記符号化ステップにおいて読み出されて、上記符号化処理が行われた静止画像情報を順次廃棄する映像バッファリングステップと、上記映像バッファリングステップにおける上記一連の静止画像情報の保存を、上記与えられたフレームレートに

基づいて決定される一定のフレームレートにおいて行うように制御するフレームレート制御ステップとを実行し、

上記映像バッファリングステップにおいて一時蓄積された上記原映像情報の蓄積量に基づいて上記判断を行うものであることを特徴とする映像符号化方法。

【請求項7】 音声に対して、帯域分割符号化方式により符号化を行う音声符号化方法において、符号化処理に用いる数値である、設定周波数 f_s と、変換定数 n とを記憶する記憶ステップと、符号化の対象である音声を入力する音声入力ステップと、

上記記憶した設定周波数 f_s に基づいて決定されるサンプリング周波数を用いて、サンプリング音声データを作成する入力音声サンプリングステップと、

上記設定周波数 f_s をサンプリング周波数として用いた場合に得られるサンプリング音声データの個数を m 個とし、上記変換定数 n に基づいて定められる数を m' として、 m' 個のサンプリング音声データを含む、 m 個の音声データからなる変換音声データを出力する音声データ変換ステップと、

上記変換音声データを、帯域分割して M 個の帯域信号を得る帯域分割ステップと、

上記記憶した設定周波数 f_s と変換定数 n とから得られる周波数 $f_s/2n$ を制限周波数として、上記帯域信号のうち、制限周波数以下の帯域信号にのみ符号化ビットを割り当てる符号化ビット割り当てステップと、

上記割り当てた符号化ビットに基づいて量子化を行う量子化ステップと、

上記量子化したデータを符号化データとして出力する符号化ステップと、

上記出力される符号化データを記録する符号化データ記録ステップとを実行することを特徴とする音声符号化方法。

【請求項8】 請求項7に記載の音声符号化方法において、

上記入力音声サンプリングステップでは、上記記憶した設定周波数 f_s をサンプリング周波数として、上記入力された音声のサンプリング処理により、 m 個のサンプリング音声データを作成するものであり、

上記音声データ変換ステップでは、上記 m 個のサンプリング音声データより、 $(n-1)$ 個おきにサンプリング音声データを抽出し、2つの隣接する上記抽出したサンプリング音声データの間に、 $(n-1)$ 個の音声データを挿入して、 m 個の変換音声データに変換するものであることを特徴とする音声符号化方法。

【請求項9】 請求項8に記載の音声符号化方法において、

上記音声データ変換ステップでは、上記抽出したサンプリング音声データがそれぞれ n 個ずつ連続する変換音声

データを作成するものであることを特徴とする音声符号化方法。

【請求項10】 請求項7に記載の音声符号化方法において、

上記入力音声サンプリングステップでは、上記記憶した設定周波数 f_s と変換定数 n とから得られる周波数 f_s/n をサンプリング周波数として、上記入力された音声のサンプリング処理により、 m/n 個のサンプリング音声データを作成するものであり、

上記音声データ変換ステップでは、上記サンプリング音声データに基づき、2つの隣接するサンプリング音声データの間に $(n-1)$ 個の音声データを挿入して、 m 個の変換音声データに変換するものであることを特徴とする音声符号化方法。

【請求項11】 請求項10に記載の音声符号化方法において、

上記音声データ変換ステップでは、上記 m/n 個のサンプリング音声データが、それぞれ n 個ずつ連続する変換音声データを作成するものであることを特徴とする音声符号化方法。

【請求項12】 請求項7ないし11のいずれかに記載の音声符号化方法において、

上記サンプリング音声データを、入力バッファに一時的に保持する音声バッファリングステップと、

上記入力バッファのデータ量を調べて、これを予め設定した値と比較し、上記比較の結果に基づいて、上記レジスタに記憶された上記変換定数 n の値を変更する入力バッファ監視ステップとを実行し、

上記入力音声サンプリングステップでは、上記サンプリング音声データを上記入力バッファに書き込むものであり、

上記音声データ変換ステップでは、上記入力バッファよりサンプリング音声データを読み出して、これを上記変換するものであることを特徴とする音声符号化方法。

【請求項13】 請求項7ないし11のいずれかに記載の音声符号化方法において、

上記符号化ステップにおいて出力される単位時間当たりの符号化データ量を調べて、これを予め設定した値と比較し、上記比較の結果に基づいて、上記レジスタに記憶された上記変換定数 n の値を変更する符号化データ監視ステップを実行するものであることを特徴とする音声符号化方法。

【請求項14】 音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化方法において、

上記符号化に用いる制御定数を記憶する制御定数記憶ステップと、

入力音声をサンプリング処理して、サンプリングデータを出力するサンプリングステップと、

上記サンプリングステップで得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する

帯域分割ステップと、

上記帯域分割ステップで得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当てステップと、

上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化ステップと、

上記量子化ステップで得られた量子化値に基づき、符号化データを出力する符号化ステップと、

10 上記記憶した制御定数に基づいて、上記帯域分割ステップ、上記符号化ビット割り当てステップ、上記量子化ステップ、および上記符号化ステップにおけるデータ処理を制御する符号化処理制御ステップとを実行することを特徴とする音声符号化方法。

【請求項15】 請求項14に記載の音声符号化方法において、

上記制御定数記憶ステップでは、

上記制御定数として、単位期間判定定数 k を単位期間判定定数レジスタに記憶するものであり、

20 上記符号化処理制御ステップは、

上記帯域分割ステップでの1回の帯域分割処理で対象とするサンプリングデータ数を p とし、 p 個のサンプリングデータに相当する時間を単位期間として、上記出力されるサンプリングデータの p 個ごとに、相当する単位期間が符号化対象期間であるか符号化対象外期間であるかの判定を、上記記憶した単位期間判定定数に基づいて行い、

上記単位期間が上記符号化対象期間と判定されたときのみ、該単位期間のサンプリングデータが上記帯域分割ステップに出力されるよう制御し、

30 上記単位期間が上記符号化対象外期間と判定されたときは、上記符号化ステップにおいて、予め記憶した固定的符号化データを符号化データとして出力するよう制御する判定制御ステップであることを特徴とする音声符号化方法。

【請求項16】 請求項15に記載の音声符号化方法において、

上記判定制御ステップでは、 i 番目の単位期間を t_i として、上記記憶した単位期間判定定数 k と任意の整数 n とから $i = n \times k + 1$ が成立するとき、上記単位期間 t_i が上記符号化対象期間であると判定するものであることを特徴とする音声符号化方法。

【請求項17】 請求項14に記載の音声符号化方法において、

上記制御定数記憶ステップでは、上記制御定数として、演算処理判定定数 q を演算処理判定定数レジスタに記憶するものであり、

上記符号化処理制御ステップは、

50 上記帯域分割ステップに内包され、上記記憶した演算処理判定定数 q に基づいて、上記帯域分割ステップにおけ

る演算処理を途中で打ち切るように制御する演算処理中止ステップであることを特徴とする音声符号化方法。

【請求項 18】 請求項 17 に記載の音声符号化方法において、

上記演算処理中止ステップでは、上記帯域分割ステップにおける基本低域通過フィルタの演算処理を、該フィルタの両端ステップ分については途中で打ち切るように制御するものであることを特徴とする音声符号化方法。

【請求項 19】 請求項 14 に記載の音声符号化方法において、

上記制御定数記憶ステップでは、上記制御定数として、帯域選択定数 r を帯域選択定数レジスタに記憶するものであり、

上記符号化処理制御ステップは、

上記帯域分割ステップが出力する帯域信号データのうち、上記記憶した帯域選択定数 r に基づいて選択したもののみに対して、上記符号化ビット割り当てステップと上記量子化ステップとにおける処理を実行するよう制御する帯域間引きステップであることを特徴とする音声符号化方法。

【請求項 20】 請求項 19 に記載の音声符号化方法において、

上記帯域間引きステップでは、上記帯域分割ステップで得られた M 個の帯域信号データ出力から、上記記憶した帯域選択定数である r 個おきに帯域信号データを選択するものであることを特徴とする音声符号化方法。

【請求項 21】 請求項 14 ないし 20 のいずれかに記載の音声符号化方法において、

音声符号化におけるデータ処理の状況を取得し、該取得した状況に応じて、上記記憶した上記制御定数の値を変更する処理状況監視ステップを実行するものであることを特徴とする音声符号化方法。

【請求項 22】 請求項 21 に記載の音声符号化方法において、

上記処理状況監視ステップでは、

サンプリングデータを入力バッファに一時蓄積する音声バッファリングステップと、

上記入力バッファに保持されるデータの量を予め設定した値と比較し、上記比較の結果に基づいて上記制御定数変更を行う入力監視ステップとを実行するものであることを特徴とする音声符号化方法。

【請求項 23】 請求項 21 に記載の音声符号化方法において、

上記処理状況監視ステップは、

上記符号化ステップにおいて単位時間あたりに出力される上記符号化データの量を、予め設定した値と比較し、上記比較の結果に基づいて上記制御定数の値を変更する符号化監視ステップであることを特徴とする音声符号化方法。

【請求項 24】 音声がデジタル化された原音声情報に

対して、帯域分割符号化方式を用いて符号化を行う音声符号化方法において、

入力音声をサンプリング処理して、サンプリングデータを出力するサンプリングステップと、

上記サンプリングステップで得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割ステップと、

上記帯域分割ステップで得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当てステップと、

上記符号化ビット割り当てステップにおける割り当てを心理聴覚分析代替制御方式により制御するビット割り当て制御ステップと、

上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化ステップと、

上記量子化ステップで得られた量子化値に基づき、符号化データを出力する符号化ステップとを実行することを特徴とする音声符号化方法。

【請求項 25】 請求項 24 に記載の音声符号化方法において、

上記ビット割り当て制御ステップは、

上記帯域分割ステップで得られた帯域信号データに対して、心理聴覚分析代替制御方式により予め定められたビット割り当て順に従って、符号化ビット割り当てを行うよう制御する順次ビット割り当てステップであることを特徴とする音声符号化方法。

【請求項 26】 請求項 24 に記載の音声符号化方法において、

上記ビット割り当て制御ステップは、

上記帯域分割ステップで得られた帯域信号データに対して、心理聴覚分析代替制御方式により予め定められた各帯域への重み付けと、各帯域信号データの有する出力レベルとに基づいた符号化ビット割り当てを行うよう制御する帯域出力適応ビット割り当てステップであることを特徴とする音声符号化方法。

【請求項 27】 請求項 24 に記載の音声符号化方法において、

上記ビット割り当て制御ステップは、

上記帯域分割ステップで得られた帯域信号データに対して、心理聴覚分析代替制御方式により予め定められた各帯域への重み付けと、各帯域毎のビット割り当て数に対する重み付けと、各帯域信号データの有する出力レベルとに基づいた符号化ビット割り当てを行うよう制御する改良型帯域出力適応ビット割り当てステップであることを特徴とする音声符号化方法。

【請求項 28】 請求項 24 に記載の音声符号化方法において、

上記ビット割り当て制御ステップは、

上記帯域分割ステップで得られた帯域信号データに対し

て、帯域信号データごとに最小可聴限界値との比較を行い、上記比較により最小可聴限界未満と判定された帯域信号データにはビット割り当てを行わず、他の帯域に対してのビット割り当てを増加するよう制御する最小可聴限界比較ステップであることを特徴とする音声符号化方法。

【請求項29】 映像と音声とを符号化するにあたり、上記2つの符号化処理に含まれる処理過程の一部または全部を、共通の計算機資源を用いて実行する映像音声符号化方法において、

単位時間毎の静止画像を表す複数の静止画像情報からなる原映像情報と、音声を表す原音声情報とから構成される映像音声情報が入力されたとき、上記原音声情報を一時的に蓄積する音声バッファリングステップと、

上記音声バッファリングステップにおいて蓄積された原音声情報を読み出し、この読み出した上記原音声情報を符号化処理し、符号化音声情報を出力する音声符号化ステップと、

映像符号化の負荷程度を表す符号化負荷基準情報を用いて、当該映像音声符号化処理についての処理能力を判断し、その判断の結果に基づいて、後述する映像符号化ステップにおける原映像情報に対する符号化を制御する符号化負荷評価ステップと、

上記符号化負荷評価ステップにおける制御に従って、入力された上記原映像情報を構成する静止画像情報を符号化処理し、符号化映像情報を出力する映像符号化ステップとを実行することを特徴とする映像音声符号化方法。

【請求項30】 請求項29に記載の映像音声符号化方法において、

上記符号化負荷評価ステップは、

上記原映像情報を構成する静止画像情報が入力されたとき、上記音声バッファリングステップにおいて蓄積された原音声情報の総量と、上記符号化負荷基準情報とに基づいて符号化負荷評価情報を求め、上記符号化負荷評価情報を予め設定された負荷限度と比較して、上記符号化負荷評価情報が上記負荷限度に達していない場合に静止画像情報を出力し、上記符号化負荷評価情報が上記負荷限度に達した場合に、上記静止画像情報を破棄するものであることを特徴とする映像音声符号化方法。

【請求項31】 請求項29に記載の映像音声符号化方法において、

アナログ映像情報を入力し、後述する映像解像度情報が出力されたとき、上記アナログ映像情報を複数の離散的デジタル画素情報からなり、上記映像解像度情報に従う解像度を持つ複数の静止画像情報で構成される原映像情報に変換し、上記映像符号化ステップにおいて処理されるよう出力する映像キャプチャステップを実行するものであり、

上記符号化負荷評価ステップでは、

上記音声バッファリングステップにおいて蓄積された原

音声情報の総量と、映像符号化の負荷程度を表す符号化負荷基準情報とに基づいて符号化負荷評価情報を求め、上記符号化負荷評価情報に基づいて、映像符号化に用いる映像の解像度を表す映像解像度情報を求め、上記映像解像度情報を出力するものであり、

上記映像符号化ステップでは、

上記映像解像度情報が出力されたとき、上記映像解像度情報に従って上記静止画像情報に対して符号化処理を行い、符号化映像情報を出力するものであることを特徴とする映像音声符号化方法。

10

【請求項32】 請求項29に記載の映像音声符号化方法において、

上記符号化負荷評価ステップでは、符号化負荷評価情報を上記映像符号化ステップにおいて処理されるよう出力するものであり、

上記映像符号化ステップでは、上記静止画像情報に対して、上記出力された符号化負荷評価情報を用いて計算される処理量だけ符号化処理を行い、符号化映像情報として出力するものであることを特徴とする映像音声符号化方法。

20

【請求項33】 請求項29ないし32のいずれかに記載の映像音声符号化方法において、

上記音声符号化ステップでは、上記音声バッファリングステップにおいて蓄積された原音声情報を読み出し、この読み出した上記原音声情報の総量を計算して処理済み音声情報量として出力し、その後、上記原音声情報を符号化処理して符号化音声情報として出力するものであり、

上記符号化負荷評価ステップでは、経過時間と、上記原音声情報の時間当たりの入力量に基づいて原音声入力量を求め、この原音声入力量と上記処理済み音声情報量との差である予測音声バッファ量を求め、上記予測音声バッファ量を用いて、上記符号化負荷評価情報を求めるものであることを特徴とする映像音声符号化方法。

30

【請求項34】 請求項29ないし32のいずれかに記載の映像音声符号化方法において、

上記符号化負荷評価ステップでは、上記静止画像情報が入力されたとき、経過時間と、上記原音声情報の時間当たりの入力量とに基づいて原音声入力量を求め、かつ、上記音声符号化ステップにおいて出力された符号化音声情報の総量に基づいて処理済み音声情報量を求め、さらに、上記求めた原音声入力量と上記求めた処理済み音声情報量との差である予測音声バッファ量を求めた後、上記予測音声バッファ量を用いて、上記符号化負荷評価情報を求めるものであることを特徴とする映像音声符号化方法。

40

【請求項35】 請求項29ないし34のいずれかに記載の映像音声符号化方法において、

上記符号化負荷評価ステップにおける、上記判断の結果の変動を監視し、上記変動に対応して、上記符号化負荷

50

基準情報を設定することを特徴とする映像音声符号化方法。

【請求項36】 映像を符号化する映像符号化装置において、

映像がデジタル化された、複数の静止画像情報からなる原映像情報に対して、上記静止画像情報の1つまたは複数を、後述する符号化パラメータに従って符号化する映像符号化手段と、

1つ以上の解像度を一の符号化パラメータとし、フレーム内符号化、順方向予測符号化、逆方向予測符号化、及び双方向予測符号化の各タイプを含む符号化タイプのうち1つ以上の符号化タイプを他の符号化パラメータとして、上記符号化手段の処理量を決定するものである符号化パラメータを、与えられたフレームレートに基づいて決定する符号化パラメータ決定手段とを備えたことを特徴とする映像符号化装置。

【請求項37】 音声に対して、帯域分割符号化方式により符号化を行う音声符号化装置において、符号化処理に用いる数値である、設定周波数 f_s と、変換定数 n とを記憶するレジスタと、符号化の対象である音声を入力する音声入力手段と、上記記憶した設定周波数 f_s に基づいて決定されるサンプリング周波数を用いて、サンプリング音声データを作成する入力音声サンプリング手段と、上記設定周波数 f_s をサンプリング周波数として用いた場合に得られるサンプリング音声データの個数を m 個とし、上記変換定数 n に基づいて定められる数を m' として、 m' 個のサンプリング音声データを含む、 m 個の音声データからなる変換音声データを出力する音声データ変換手段と、

上記変換音声データを、帯域分割して M 個の帯域信号を得る帯域分割手段と、

上記記憶した設定周波数 f_s と変換定数 n とから得られる周波数 $f_s/2n$ を制限周波数として、上記帯域信号のうち、制限周波数以下の帯域信号にのみ符号化ビットを割り当てる符号化ビット割り当て手段と、

上記割り当てた符号化ビットに基づいて量子化を行う量子化手段と、

上記量子化したデータを符号化データとして出力する符号化手段と、

上記出力される符号化データを記録する符号化データ記録手段とを備えたことを特徴とする音声符号化装置。

【請求項38】 音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化装置において、上記符号化に用いる制御定数を記憶する制御定数記憶手段と、

入力音声をサンプリング処理して、サンプリングデータを出力するサンプリング手段と、

上記サンプリング手段で得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域

分割手段と、

上記帯域分割手段で得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当て手段と、

上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化手段と、

上記量子化手段で得られた量子化値に基づき、符号化データを出力する符号化手段と、

10 上記記憶した制御定数に基づいて、上記帯域分割手段、上記符号化ビット割り当て手段、上記量子化手段、および上記符号化手段におけるデータ処理を制御する符号化処理制御手段とを備えたことを特徴とする音声符号化装置。

【請求項39】 音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化装置において、

入力音声をサンプリング処理して、サンプリングデータを出力するサンプリング手段と、

20 上記サンプリング手段で得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割手段と、

上記帯域分割手段で得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当て手段と、

上記符号化ビット割り当て手段における割り当てを心理聴覚分析代替制御方式により制御するビット割り当て制御手段と、

30 上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化手段と、

上記量子化手段で得られた量子化値に基づき、符号化データを出力する符号化手段とを備えたことを特徴とする音声符号化装置。

【請求項40】 映像と音声とを符号化するにあたり、上記2つの符号化処理に含まれる処理過程の一部または全部を、共通の計算機資源を用いて実行する映像音声符号化装置において、

単位時間毎の静止画像を表す複数の静止画像情報からなる原映像情報と、音声を表す原音声情報とから構成される映像音声情報が入力されたとき、上記原音声情報を一時的に蓄積する音声バッファリング手段と、

上記音声バッファリング手段において蓄積された原音声情報を読み出し、この読み出した上記原音声情報を符号化処理し、符号化音声情報を出力する音声符号化手段と、

映像符号化の負荷程度を表す符号化負荷基準情報を用いて、当該映像音声符号化装置の処理能力を判断し、その判断の結果に基づいて、後述する映像符号化手段に対しての上記原映像情報の出力を制御する符号化負荷評価手段と、

上記符号化負荷評価手段の制御に従って、上記原映像情報を構成する静止画像情報が入力されたとき、上記静止画像情報を符号化処理し、符号化映像情報を出力する映像符号化手段とを備えたことを特徴とする映像音声符号化装置。

【請求項41】 映像を符号化処理する映像符号化プログラムを記録した記録媒体において、映像がデジタル化された、複数の静止画像情報からなる原映像情報に対して、上記静止画像情報の1つまたは複数を、後述する符号化パラメータに従って符号化する映像符号化ステップと、

1つ以上の解像度を一の符号化パラメータとし、フレーム内符号化、順方向予測符号化、逆方向予測符号化、及び双方向予測符号化の各タイプを含む符号化タイプのうち1つ以上の符号化タイプを他の符号化パラメータとして、上記符号化ステップの処理量を決定するものである符号化パラメータを、与えられたフレームレートに基づいて決定する符号化パラメータ決定ステップとを実行する符号化プログラムを記録したことを特徴とする映像符号化プログラム記録媒体。

【請求項42】 音声に対して、帯域分割符号化方式により符号化を行う音声符号化プログラムを記録した記録媒体において、

符号化処理に用いる数値である、設定周波数 f_s と、変換定数 n とを記憶する記憶ステップと、符号化の対象である音声を入力する音声入力ステップと、

上記記憶した設定周波数 f_s に基づいて決定されるサンプリング周波数を用いて、サンプリング音声データを作成する入力音声サンプリングステップと、

上記設定周波数 f_s をサンプリング周波数として用いた場合に得られるサンプリング音声データの個数を m 個とし、 $m \geq m'$ である、上記変換定数 n に基づいて定められる数を m' として、 m' 個のサンプリング音声データを含む、 m 個の音声データからなる変換音声データを出力する音声データ変換ステップと、

上記変換音声データを、帯域分割して M 個の帯域信号を得る帯域分割ステップと、

上記記憶した設定周波数 f_s と変換定数 n とから得られる周波数 $f_s/2n$ を制限周波数として、上記帯域信号のうち、制限周波数以下の帯域信号にのみ符号化ビットを割り当てる符号化ビット割り当てステップと、上記割り当てた符号化ビットに基づいて量子化を行う量子化ステップと、

上記量子化したデータを符号化データとして出力する符号化ステップと、

上記出力される符号化データを記録する符号化データ記録ステップとを実行する符号化プログラムを記録したことを特徴とする音声符号化プログラム記録媒体。

【請求項43】 音声に対して、帯域分割符号化方式を

用いて符号化を行う音声符号化プログラムを記録した記録媒体において、

上記符号化に用いる制御定数を記憶する制御定数記憶ステップと、

入力音声をサンプリング処理して、サンプリングデータを出力するサンプリングステップと、

上記サンプリングステップで得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割ステップと、

10 上記帯域分割ステップで得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当てステップと、

上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化ステップと、

上記量子化ステップで得られた量子化値に基づき、符号化データを出力する符号化ステップと、

20 上記記憶した制御定数に基づいて、上記帯域分割ステップ、上記符号化ビット割り当てステップ、上記量子化ステップ、および上記符号化ステップにおけるデータ処理を制御する符号化処理制御ステップとを実行する符号化プログラムを記録したことを特徴とする音声符号化プログラム記録媒体。

【請求項44】 音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化プログラムを記録した記録媒体において、

入力音声をサンプリング処理して、サンプリングデータを出力するサンプリングステップと、

30 上記サンプリングステップで得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割ステップと、

上記帯域分割ステップで得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当てステップと、

上記符号化ビット割り当てステップにおける割り当てを心理聴覚分析代替制御方式により制御するビット割り当て制御ステップと、

40 上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化ステップと、

上記量子化ステップで得られた量子化値に基づき、符号化データを出力する符号化ステップとを実行する符号化プログラムを記録したことを特徴とする音声符号化プログラム記録媒体。

【請求項45】 映像と音声とを符号化するにあたり、上記2つの符号化処理に含まれる処理過程の一部または全部を、共通の計算機資源を用いて実行する映像音声符号化プログラムを記録した記録媒体において、

単位時間毎の静止画像を表す複数の静止画像情報からなる原映像情報と、音声を表す原音声情報とから構成され

る映像音声情報が入力されたとき、上記原音声情報を一時的に蓄積する音声バッファリングステップと、上記音声バッファリングステップにおいて蓄積された原音声情報を読み出し、この読み出した上記原音声情報を符号化処理し、符号化音声情報を出力する音声符号化ステップと、

映像符号化の負荷程度を表す符号化負荷基準情報を用いて、当該映像音声符号化処理についての処理能力を判断し、その判断の結果に基づいて、後述する映像符号化ステップにおける原映像情報に対する符号化を制御する符号化負荷評価ステップと、

上記符号化負荷評価ステップにおける制御に従って、入力された上記原映像情報を構成する静止画像情報を符号化処理し、符号化映像情報を出力する映像符号化ステップとを実行する符号化プログラムを記録したことを特徴とする映像音声符号化プログラム記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、映像・音声符号化方法、符号化装置、及び符号化プログラム記録媒体に関し、特に、汎用計算機資源を用いてソフトウェア制御により、映像、音声、又は映像音声を取り込みに伴って符号化する符号化方法、符号化装置、及び符号化プログラム記録媒体に関する。

【0002】

【従来の技術】本来アナログのデータである映像や音声をデジタル化し、デジタルの映像データや音声データを得る技術については、デジタルデータの記録、伝送、編集、複製および伝送等における扱いの容易さから普及と発展が著しい分野となっている。デジタル化の利点の一つとして、データを容易に圧縮可能であるという点が挙げられ、特に記録や伝送のためには圧縮符号化は重要な技術である。かかる圧縮符号化の技術については、国際的規格も確立され、中でもMPEG規格は、映像や音声を扱い得る一般的なデジタル規格として普及している。

【0003】また、近年、コンピュータならびにVLSI等の半導体デバイスの高速化、低価格化に伴い、マルチメディア対応パソコンと呼ばれるパソコンが低価格で市場に出回り、従来デコードハードウェアの追加により行っていた、圧縮符号化したデジタルデータである映像および音声の再生が、家庭用・個人用パーソナルコンピュータ上においても、ソフトウェアによって簡単に出来るようになってきている。また、これにともなって、インターネット等によっても映像や音声の配信などが行われるようになっており、MPEG等の規格に準拠した映像および音声の符号化データの利用範囲が拡大している。

【0004】一方、それらの映像や音声の符号化データを作る符号化（エンコード）処理については、一般的には、家庭用・個人用パーソナルコンピュータではソフトウェア処理が困難であり、専用ハードウェアを追加して

行うこととなる。また、一旦ファイルとして記録して後にソフトウェア処理により符号化を行うことも可能ではあるが、映像や音声の入力時間の、何倍もの処理時間をかけて変換を行うものとなることから、一般ユーザにとって魅力的なソフトウェアとは言い難い。

【0005】一般のパソコンユーザーが簡単に、動画を含む映像や音声を取り込んで符号化データを作成出来るようになるためには、キャプチャーボードやサウンドボードを用いて動画や音声を取り込み、この取り込みにともなっている、ソフトウェアによる実時間符号化処理が可能であることが望ましいものであり、ハードウェア面での発展・普及に伴い、開発の望まれる分野である。映像、音声、または映像音声符号化の分野における現状について、以下に、「A映像符号化」「B音声符号化」「C映像音声符号化」を行う、従来の技術による装置について説明する。

【0006】A. 従来の技術による映像符号化装置

静止画像もしくは動画画像を含む映像を、リアルタイムでデジタル化してコンピュータに取り込み、該取り込みに伴っての符号化処理を施すことは、動画像圧縮の国際標準であるMPEG方式を用いて映像をリアルタイムで符号化するパソコン用拡張カードなどを用いることによって実現されている。

【0007】図58は、このような専用ハードウェアを有するコンピュータにおいて実現された、映像符号化装置の構成を示すブロック図である。図示するように、従来の技術による映像符号化装置は、符号化手段5001と符号化パラメータ決定手段5002とからなり、映像を入力画像データとして入力し、映像符号化データを出力する。符号化手段5001には、DCT処理手段5003、量子化手段5004、可変長符号化手段5005、ビットストリーム生成手段5006、逆量子化手段5007、逆DCT処理手段5008、および予測画像生成手段5009が含まれている。

【0008】図において、符号化手段5001は、映像がデジタル化された、一連の静止画像からなる映像データを入力画像データとして入力し、設定された符号化パラメータに従って符号化処理し、符号化データを出力する。入力画像データを構成する個々の静止画像データをフレーム画像と呼ぶ。また、符号化パラメータは、符号化タイプと、解像度とを指示するものとして、後述する符号化パラメータ決定手段5002から与えられるものである。符号化パラメータ決定手段5002は、フレーム内符号化処理、またはフレーム間符号化処理を示す符号化タイプと、解像度とを決定して、これらを符号化パラメータとして符号化手段5001に出力する。

【0009】符号化手段5001の内部においては、入力画像データに対してまずDCT手段5003がDCT（離散コサイン変換）処理を行ってDCT変換データを出力し、次に、量子化手段5004が、DCT変換デー

タに対して量子化処理を行って量子化データを出力し、次に可変長符号化手段5005が量子化データに対して可変長符号化処理を行うことによって、圧縮符号化された可変長符号化データが作成される。可変長符号化データはビットストリーム生成手段5006に輸入され、ビットストリーム生成手段5006から、伝送や記録を行うことのできるビットストリームとして、当該映像符号化装置の符号化結果である符号化データが出力される。

【0010】逆量子化手段5007は、量子化手段5004から出力された量子化データに対して、量子化処理の逆処理である逆量子化処理を行って逆量子化データを出力し、次に逆DCT手段5008が逆量子化データに対して、DCT処理の逆処理である逆DCT処理を行って、逆DCT変換データを出力し、逆DCT変換データは予測画像生成手段5003に輸入され、予測画像データとして出力されることとなる。符号化パラメータに従って、予測画像を用いた符号化処理が行われる場合には、この予測画像データと入力画像データとの差分データがDCT手段5003に輸入されることにより、符号化手段5001においてはフレーム間符号化が行われることとなる。

【0011】このように構成された、従来の技術による映像符号化装置による、映像符号化処理の際の動作について、以下に説明する。まず、符号化処理に先立って、符号化パラメータ決定手段5002は、符号化タイプと解像度とについて符号化パラメータを決定し、これを符号化手段5001に出力する。

【0012】一般に、圧縮符号化にあたっては、1フレーム（1画面相当）の静止画像について、その空間的相関関係（フレーム内の相関関係）に基づいて、冗長性を除いて圧縮を行うフレーム内符号化と、時間的に近接する、例えば連続するフレームの静止画像について、その時間的相関関係（フレーム間の相関関係）に基づいて冗長性を除いて圧縮を行うフレーム間符号化とがある。

【0013】従来の技術による映像符号化装置では、基本的にフレーム内符号化を行うものであるが、フレーム間符号化をも行うことで、高圧縮率の符号化データが得られる。しかし、フレーム間符号化を行うためには、符号化の逆処理である復号化処理や、動き検出・動き補償処理により、予測画像を生成し、この予測画像と符号化対象画像との差分を取得するものであるため、これらの処理を必要とする分、装置にとっての処理負担の増大を招くこととなる。フレーム間符号化を行う場合の予測画像の生成については、直前に処理したデータに基づいて予測を行う順方向予測、直後に処理されるデータに基づいて予測を行う逆方向予測、および順方向、または逆方向予測を行う双方向予測のいずれかがなされる。なお、フレーム内符号化を「I」、順方向予測符号化を「P」、双方向予測符号化（逆方向を含む）を「B」と表記する。

【0014】また、画像の解像度については、「320×240」、または「160×120」など1画面あたりの縦横の画素（ピクセル）数で表わされることが一般的であり、高解像度において、すなわち1画面に多くの画素を有することとして処理を行う方が、再生画質の良好なデータが得られるものであるが、それだけ処理対象が増えることから、処理負担の増大を招くこととなる。

【0015】又、MPEG規格に準拠するためには、一定の転送レートにおいて、データの入出力や転送が行われる必要があり、符号化処理においては、この転送レートを満足するように処理が行われ、符号化データが出力される必要がある。映像を処理対象とする場合には、転送レートは、1秒当たりのフレーム数で表わされるフレームレートで表示することが一般的である。

【0016】従って、当該映像符号化装置の処理能力を考慮して、このフレームレートを満足しつつ、映像取り込みに伴うリアルタイム処理をなし、かつなるべく高圧縮率で、再生画質の良い（高解像度の）符号化データが得られるように、符号化パラメータが設定されることが望ましい。

【0017】従来の技術による映像符号化装置においては、符号化パラメータはこれらの要因を考慮して予め設定されたものとして、符号化パラメータ手段5002はこの設定されたパラメータを保持し、符号化の際には符号化手段5001に出力するものとすることができる。また、符号化パラメータのうち、符号化タイプに関しては、シーンチェンジ等の入力映像の情報に基づいて、符号化タイプを決定する方法が、「画像符号化装置（特開平8-98185号公報）」において、開示されている。

【0018】符号化手段5001に輸入された符号化パラメータのうち、解像度を示すパラメータはDCT処理手段5003に輸入され、処理に用いられる。また、符号化タイプを示すパラメータは、DCT処理手段5003に対する入力を、入力画像データそのものとするか、予測画像との差分にするかの切り替えの制御に用いられることとなる。

【0019】DCT処理手段5003は、符号化パラメータ決定手段5002から輸入された解像度に基づき、入力されたフレーム画像、または差分データに対して、DCT処理をして、DCT変換データを出力する。DCT処理は、一般には、対象となるデータを8画素×8画素のブロックに分割し、分割したブロックごとに2次元離散コサイン変換することが行われる。次いで、量子化手段104は、DCT変換データに対して、ある定められた値を用いて量子化処理を行い、量子化データを出力する。量子化処理は、一般には、量子化ステップの値（上記の定められた値）を用いた除算処理により行われる。そして、可変長符号化手段105は、量子化データを可変長符号化し、可変長符号化データを出力する。可

変長符号化は、符号化処理におけるビット割当において、頻度の高いものに対してビット数の少ない符号を割り当てることにより、全体のデータ量を小さくするものである。ビットストリーム生成手段106は、可変長符号化手段105が出力した可変長符号化データより、当該映像符号化装置の装置出力として、符号化結果であるビットストリームを生成して出力する。

【0020】フレーム間符号化が実行される場合には、次の動作が行われる。逆量子化手段5007は、量子化手段5004が出力する量子化データを逆量子化し、逆量子化データを出力する。次いで、逆DCT処理手段5008が、逆量子化データに対して、DCT処理手段5003が分割した8画素×8画素のブロックごとに、2次元離散コサイン変換の逆処理である2次元逆離散コサイン変換を実行し、逆DCT変換データを出力する。予測画像生成手段5009は、逆DCT変換データに基づいて予測画像を生成し、これを出力する。DCT処理手段5003には、入力画像データと予測画像との差分データが入力されることとなる。

【0021】B. 従来の技術による音声符号化装置
音声符号化方法については、MPEG Audio方式に準拠した帯域分割符号化方式による符号化方法が、人間の声や音楽、自然環境の音、様々な効果音など、広い帯域にわたる音声一般を符号化するのに用いられる。高性能なマルチメディアパソコン等では、標準装備されることが一般的なサウンドボードを用いて取り込んだ音声を取り込みにもなっている実時間符号化処理を行うことも可能である。従来技術による音声符号化装置の第1の例としては、上記の帯域分割符号化方式によって、入力した音声を符号化するものについて説明する。

【0022】一方、MPEG1 Audioに準拠した音声符号化の方法として、心理聴覚分析を応用する方法がある。本来、MPEG1 Audioに準拠したエンコーダでは、聴覚心理モデルを用いて人間の聴覚能力の限界や、マスキング効果を考慮した上で、各帯域にビットを割り当てる優先順位を決める。これは、人間の静的および動的な聴覚特性にあわせた高能率符号化を行うためのものであるが、MPEG1 Audio規格のデータフォーマットには影響せず、たとえこれを行わなくとも、MPEG1 Audio符号化データは作成可能である。また、後述するように、心理聴覚分析の処理は処理負荷が大きく、従来技術の第1の例に示したように、この処理を省くことにより、CPUに対する大幅な処理負荷軽減を図ることが可能となる。ただし、心理聴覚分析を応用しない分、再生音質は低下することとなる。従来技術による音声符号化の第2の例としては、かかる心理聴覚分析を応用した音声符号化について説明する。

【0023】B-1. 従来の技術による音声符号化装置の第1例

図59は、従来技術の第1の例による音声符号化装置の

構成を示すブロック図である。図示するように、第1の例による音声符号化装置は、音声入力部2551、入力音声サンプリング部2553、帯域分割部2555、符号化ビット割り当て部2556、量子化部2557、符号化部2558、および符号化データ記録部2559から構成されている。

【0024】図において、音声入力部2551は、符号化を行う音声を入力する。一般的には音声はマイクロホンから、あるいはライン入力として入力される。入力音声サンプリング部2553は、サウンドボードの入力機能および制御プログラムによって実現され、音声入力部2551が入力した音声に対してサンプリング処理を行う。帯域分割部2555は、サンプリング処理されたデータを帯域分割する。符号化ビット割り当て部2556は、帯域分割部2555が分割した帯域のそれぞれに対して、符号化ビットを割り当てる。量子化部2557は、符号化ビット割り当て部2556の割り当てた符号化ビット数に従って、量子化処理を行う。符号化部2558は、量子化部2557の出力する量子化値を符号化音声データとして出力する。2555～2558はいずれも、コンピュータのCPU、メインメモリ、およびプログラムで実現される。符号化データ記録部2559は、磁気記憶装置等の記憶装置、および該記憶装置の制御プログラムで実現され、出力された符号化データを記録する。

【0025】図60は、従来の音声符号化方法のフローチャート図、図61はサンプリング処理を、図62～図63は帯域分割を説明するための図である。以下に、図59～図63を参照し、図60のフローに従って、従来技術の第1の例による音声符号化装置の動作を説明する。

【0026】図60のステップ1では、入力音声サンプリング部2553が、予め設定したサンプリング周波数 f_s で入力音声信号をサンプリングしてサンプリングデータとする。図61に示すように、入力音声は時間と音圧の関係を示すグラフとして表現される。サンプリングはこの入力音声をサンプリング周期といわれる時間 t_s ごとに等分するように行われるが、このサンプリング周期 t_s と上記サンプリング周波数 f_s との間には、図示したように逆数関係が成立している。

【0027】図60のステップ2以降は、CPUの制御によりソフトウェア的に行われる演算処理を中心とする処理である。ステップ2において、サンプリングデータは、帯域分割部2555によりM個の周波数帯域に帯域分割される。図62は音声データを全帯域入力信号として、これを12の帯域に分割する場合を示す概念図であり、帯域0信号BPF0から帯域11信号BPF11までの12個の帯域信号が作成されることを示している。図63はこのように12個に帯域分割された帯域信号を示す図である。この図において帯域分割された信号は、

図61と異なり、音圧を時間ではなく周波数との関係において表したものである。

【0028】MPEGオーディオの場合、レイヤ1～3が規定されており、1→2→3の順に再生音質が良好なものとなるが、必要なハードウェア性能は高くなり、ハードウェア規模が増大する。ここで、レイヤ1に適応した音声符号化を行う場合には、1回の帯域分割処理で対象とする入力音声サンプル数 p は $p=32$ となる。そして入力音声サンプルとしては対象とする32サンプルを中心に前後512サンプルを用いて、32帯域へ分割し

帯域ごとの音声データを出力する。

【0029】ステップ2で帯域分割によって得られたM個の帯域信号データは、帯域分割部2555から量子化部2557に渡される。一方、ステップ3で符号化ビット割り当て部2556は、M個の帯域信号の全てに対して、符号化ビットを割り当てる。そしてステップ4では、量子化部2557が、帯域分割部2555から渡された帯域信号データに対して、符号化ビット割り当て部2556により割り当てられた符号化ビット数に応じて、帯域ごとに量子化を行って量子化値とする。次いで、ステップ5では符号化部2558が、その量子化値を符号化して出力し、符号化データ記録部2559によって、出力された符号化データが記録される。

【0030】音声入力が続く間は、ステップ1～5が繰り返し実行されることにより、音声が入力され続け、実時間符号化処理が行われて、符号化データが出力され、記録され続ける。音声入力が終わると、速やかに符号化処理は終了する。記憶装置に格納された符号化データは、MPEG再生可能なデータとして保存される。あるいは記録格納する代わりに、符号化データがネットワーク等によって伝送され利用されることも可能である。以上が、従来技術の第1の例として示した、音声取り込みに伴って実時間で符号化データを得る音声符号化装置である。

【0031】B-2. 従来技術による音声符号化装置の第2例

図64は、従来技術の第2の例の音声符号化装置の構成を示すブロック図である。図示するように、第2の例による音声符号化装置は、音声入力部2651、入力音声サンプリング部2653、帯域分割部2655、量子化部2657、符号化部2658、符号化データ記録部2659、FFT部2660、心理聴覚分析部2661、および符号化ビット割り当て部2662から構成されている。この装置は第1の例の装置にFFT部2660と心理聴覚分析部2661を追加した構成である。

【0032】図において、FFT（高速フーリエ変換）部2660は、信号に対してフーリエ変換処理を施して、心理聴覚分析を行えるようにする。心理聴覚分析部2661は、FFT部2660において処理された信号に対し、最小可聴限界との比較や、マスキング効果の分

析を行う。符号化ビット割り当て部2662は、心理聴覚分析部2661の分析結果に基づいて、人間の耳に聞こえる信号に対する符号化ビットの割り当てを相対的に増やすように、符号化ビットの割り当てを行う。音声入力部2651、入力音声サンプリング部2653、帯域分割部2655、量子化部2657、符号化部2658、および符号化データ記録部2659は、第1の例の2551～2555、および2557～2559と同様であるので、説明を省略する。

【0033】図65は、MPEG1Audio符号化のフローチャート、図66は最小可聴限界を示す図である。以下に第2の例の音声符号化装置の動作について、図64～図66を参照して説明する。図65のフローにおけるステップ1からステップ2は第1の例と同様に行われ、M個の帯域に分割された信号が得られる。一般的な例として、 $M=32$ 個の帯域信号を得るものとする。第1例と同様に、帯域信号は帯域分割部55から量子化部57に渡される。

【0034】一方ステップ3ではFFT部2660が、サンプリングした入力音声データを、高速フーリエ変換（FFT）処理によってL個の帯域に分割した後、心理聴覚分析部2661に渡し、心理聴覚分析部2661はこのL個の信号に対して分析を行う。例えばMPEGオーディオのレイヤ1に従った場合は、512個のサンプリングデータが用いられるが、FFT部2660では高速フーリエ変換処理によって $L=256$ 個の帯域に分割がなされる。なお、レイヤ2の場合には、1024サンプルを用いて512帯域の出力を行うので、それだけ処理負担は増大する。

【0035】心理聴覚分析部2661は、各帯域信号について、図66に示すような人間の耳に聞こえない限界レベルである最小可聴限界との比較を行う。なお、図66は32帯域への分割として示してあるが、この例のように256個などと分割数が増加した場合でも最小可聴限界のグラフは同じものであって、図66に示したのと同じ範囲に対して横軸（帯域）についての細分化が行われることとなる。

【0036】心理聴覚分析部2661の分析により、最小可聴限界未満とされた帯域に対しては、後段の処理においてビット割り当てが行われず、その分、他の帯域により多くのビットが割り当てられることとなる。

【0037】又、人間の聴覚については、比較的小さな音、すなわち音圧の小さな信号は、周波数的に、または時間的に近接する大きな音、即ち音圧の大きい信号があるときには、聞き取られないというマスキング現象があることが認められている。そこで心理聴覚分析部2661は、各帯域の信号について近接する信号との関係を調べ、マスキング現象によってマスクされる（聞き取れない）信号を検出する。ここで検出された信号についても、後段の処理においてビットが割り当てられず、その

分、他の帯域により多くのビットが割り当てられることとなる。

【0038】図65のフローのステップ5では、符号化ビット割り当て部2657が、心理聴覚分析部2661の分析結果に応じて符号化ビットの割り当てを行う。ここでは、L帯域の分析結果について、M帯域に対する割り当てが行われる。従って、人間の耳に聞こえない、または聞こえにくい信号についてはビットが割り当てられず、その分、よく聞こえる信号に対して多くのビットが割り当てられる。ステップ6以降については、第1の例と同様であり、ステップ1〜7が繰り返されることで、音声入力に伴った音声符号化が行われる。

【0039】このようにして、人間の聴覚により聞き取りやすい音声に、より多くの符号化ビットが割り当てられることにより、心理聴覚分析を取り入れたMPEGオーディオの音声符号化では、再生音質の良好な符号化音声データを得ることが可能となる。

【0040】C. 従来の技術による映像音声符号化装置図67は従来の技術による映像音声符号化装置の概略構成を示す図である。図示するように、従来の技術による映像符号化装置は、ビデオカメラ2701、音声キャプチャ部2702、音声符号化部2703、映像キャプチャ部2704、および映像符号化部2705から構成されている。当該映像音声符号化装置からは、図示するように符号化音声情報と符号化映像情報とが装置出力として出力され、これらは必要に応じて伝送されたり記録されることとなる。

【0041】同図において、ビデオカメラ2701は、映像音声情報を取り込み、アナログ音声情報とアナログ映像情報とに分けて出力する。音声キャプチャ部2702は、ビデオカメラ2701から出力されたアナログ音声情報を入力し、離散的なデジタルデータからなるデジタル原音声情報として出力する。音声符号化部2703は、原音声情報を圧縮符号化処理し、符号化音声情報を出力する。映像キャプチャ部2704は、ビデオカメラ2701から出力されたアナログ映像情報を入力し、離散的なデジタルデータからなり、単位時間ごとの静止画像の複数枚から構成されるデジタルの原映像情報を出力する。映像符号化部2705は、映像キャプチャ部2704から出力された原映像情報を入力し、圧縮符号化して符号化映像情報を出力する。

【0042】このように構成される従来の技術による映像音声符号化装置における、映像音声のとりこみにともなう符号化の際の動作を以下に説明する。まず、ビデオカメラ2701が映像音声情報を取り込み、アナログ音声情報とアナログ映像情報とに分けて出力する。

【0043】アナログ音声情報は、音声キャプチャ部2702に入力され、音声キャプチャ部2702は、アナログ/デジタル変換処理によって、デジタル原音声情報を作成して、これを音声符号化部2703に出力する。

一方、アナログ映像情報は、映像キャプチャ部2704に入力され、映像キャプチャ部2704は、アナログ/デジタル変換処理によって、複数の静止画像情報からなるデジタル原映像情報を作成して、これを映像符号化部2705に出力する。

【0044】音声符号化部2703は、原音声情報に対して符号化処理を行って、符号化音声情報を出力する。一方、映像符号化部2705は、原映像情報に対して符号化処理を行って符号化映像情報を出力する。

【0045】ビデオカメラ2701から、映像音声のとりこみが続く間は、音声キャプチャ部2702、音声符号化部2703、映像キャプチャ部2704、および映像符号化部2705によるデジタル化と符号化が継続され、映像音声のとりこみが終了した後、デジタル化と符号化も終了する。

【0046】

【発明が解決しようとする課題】従来の技術のA〜Cの例において示したように、従来の技術による映像符号化装置、音声符号化装置、および映像音声符号化装置は、映像、音声、または映像音声のとりこみに伴って、符号化処理を行い、符号化映像データ、符号化音声データ、または符号化映像データと符号化音声データとを出力して、記録や伝送しての利用に供するものである。

【0047】A. 従来の技術による映像符号化の問題点しかしながら、従来技術のAに示す、リアルタイム処理の可能な映像符号化装置を、例えばパーソナルコンピュータ(PC)等の汎用計算機システムにおいて、符号化処理を行うソフトウェアを実行するものとして実現しようとした場合には、当該ソフトウェアは様々な環境(周辺機器、ネットワーク環境等)におかれた、多様な性能のハードウェアにおいて実行され得るものであることから、以下のような問題点につながるものとなる。

【0048】例えば上記のリアルタイム映像符号化装置をPC上で動作するアプリケーションソフトとして実現するものとして、入力された映像を、とりこみに伴ってのリアルタイム処理をし、320×240の解像度においてMPEG1の規格に従った符号化をする場合に、符号化タイプとしてフレーム内符号化である「I」と、フレーム間予測符号化である「P」および「B」について「IBBPBB」の順番で繰り返すパターンを選択したとする。この場合に、基本的なハードウェア性能が比較的高い場合、例えば動作周波数166MHzの制御装置(CPU)を有する場合に、かかるソフトウェア処理を行うのであれば、上記の設定に従い、「IBBPBB」のパターンを有する符号化タイプにおいて6つのフレーム画像を処理して、6/30秒を要するものとする。この場合であれば、結果として30フレーム/秒で映像をリアルタイム符号化できることとなる。

【0049】一方、基本的なハードウェア性能が低い場合、例えば動作周波数100MHzの制御装置(プロセッ

サ、CPU)を有する場合に、かかるソフトウェア処理を行うものであれば、上記のような符号化処理を6/30秒では行い得なくなり、得られた符号化データのフレームレートが小さなものとなる。符号化結果において、フレームレートが、30(フレーム/秒)以下であると、その符号化結果からの再生により得られる映像は動きがぎこちないものになるので、このような場合には良好な符号化をえないことになってしまう。

【0050】同様の事態は、例えばこのようなソフトウェア処理を、マルチタスクオペレーティングシステム上の一つのタスク(作業)として実行する際に、他のタスク(作業)として、ワードプロセッサ等の別のアプリケーションソフトが実行された場合や、割り込みによる中断があった場合などでは、比較的高性能なハードウェア環境においても起こり得るものである。

【0051】また、同様の符号化を解像度「320×240」でならば問題なく実行できても、解像度「640×400」で実行したならば、処理速度が十分なものとならず、フレームレート低下による問題が起こることでもある。

【0052】以上は、ハードウェア性能が不足となる場合の問題点であるが、逆に高性能なハードウェアを生かし切れないこととなる事態も起こり得る。例えば、動作周波数166MHzの制御装置(CPU)を有するハードウェアにおいて、入力された映像を、とりこみに伴ってのリアルタイム処理をし、320×240の解像度においてMPEG1の規格に従った符号化をする場合に、符号化タイプとしてフレーム内符号化である「I」のみを用いることとして処理するならば、この条件での処理であれば、「I」タイプにおいて1フレーム画像を1/30秒で処理できることから、結果として30フレーム/秒で映像をリアルタイム符号化できるものとなる。

【0053】これに対して、基本的なハードウェア性能がより高い場合、例えば動作周波数200MHzの制御装置(CPU)を有する場合に、かかるソフトウェア処理を行うものであれば、本来上記「I」タイプの1フレーム画像処理を1/30秒より短時間で実行できるものであることから、ハードウェア性能を生かせない事態となってしまう。高性能な制御装置を用いることは、それだけコストも高いものとなることから、このことは、映像符号化装置としてはコストパフォーマンスが良くないということを意味する。

【0054】この場合には、例えば、「I」タイプのみならず、「P」や「B」の符号化タイプを用いて処理を行えば、同等の画質において圧縮率の高い符号化データが得られるものであるから、上記のように「I」タイプのみにおいて圧縮率の低い符号化データを生成することは、結局装置資源を活用していなかったということになる。

【0055】同様のことは、マルチタスクオペレーティ

ングシステム上での実行などで、予想以上に計算機資源(CPU時間の割り当て)を利用できた場合や、「320×240」よりも低解像度である「160×120」の解像度で符号化を行う場合などにも起こり得るものである。

【0056】B. 従来の技術による音声符号化の問題点 B-1の第1の例のように行われる従来の音声符号化方法では、サウンドボードを装備したマルチメディアパソコン等において、ソフトウェア処理により、音声取り込みに伴ったリアルタイムでの音声符号化が可能である。しかし、このことは即ち、音声の入力に伴った実時間符号化に十分な性能を有する装置が用いられることが前提であり、目的に適應して設計されたLSIで構成したり、十分な性能を持つ制御装置(プロセッサ)を選択したりして対応がされていた。または、十分な性能を有していないプロセッサでの処理においては、途中でデータをファイルとして記録し、この記録されたデータを処理するなど、実時間の何倍かの時間を許して符号化処理を行うほかはなかった。

【0057】すなわち、MPEGAudioなどで用いられる帯域分割符号化処理を、ソフトウェア的にCPUで実行させて、音声入力に伴った実時間処理を行おうとする場合、当該ソフトウェアを実行するハードウェア環境、代表的なものとしてはCPU性能によって可能か、不可能かが決定されてしまい、例えば、CPU性能に対応した符号化レベルで実時間符号化することなどは出来なかった。

【0058】又、上記のように構成された音声符号化装置は、当初の設定に従って、一定レートで音声を入力し、実時間符号化処理を一定レートで行うように設計されているものであるが、汎用のパーソナルコンピュータ等であると、マルチタスク処理による他のタスクの影響、その他の割り込み発生等で、CPUの処理能力が割かれ、かかる状態では当初の設定に従って音声符号化処理を行えなくなることがあり、これに対応することが困難である。

【0059】次に、B-2の第2の例に示すように心理聴覚分析を行った帯域分割符号化処理では、上記のように人間の聴覚特性に応じたビット割り当てがなされることにより、再生音質の良好な符号化データが得られる。

【0060】しかし、多数の帯域への分割と、それら分割された信号についての変換処理、および比較処理の処理負担は大きく、かかる心理聴覚分析を行うことにより、一般的に約2倍程度に処理負担が増加することとなる。従って、標準的なパーソナルコンピュータレベルでは、心理聴覚分析を取り入れると、音声取り込みに伴った実時間処理を行うことは困難であり、専用のプロセッサやボード等の高性能なハードウェアの追加に頼るか、実時間処理をあきらめ、ファイルとして記録した後に時間をかけて符号化処理を行うかをせざるを得なかった。

【0061】C. 従来の技術による映像音声符号化の問題点

上述のように、上記従来の映像音声符号化装置は、音声、および映像の符号化に際しては、原音声情報（デジタル音声情報）、および原映像情報（デジタル映像情報）が直接、それぞれ対応する符号化部に入力され、それぞれに符号化処理されるものである。そのため、音声符号化部、および映像符号化部については、入力される原音声情報、および原映像情報を、例えばMPEG規格のような規格に従って、確実に処理できる能力が必要とされる。例えば、音声符号化部は、サンプリング周波数が48KHzで、1サンプル1byteの音声情報を入力するのであれば、1秒に48Kbyteの音声情報を確実に符号化できる能力が必要となる。また、映像符号化部は、横320ピクセル、縦240ピクセル、1ピクセル2byte、30fpsの映像情報が入力されるのであれば、1秒に4.6Mbyteの映像情報を確実に符号化できる能力が必要となる。

【0062】このため、従来は、音声符号化部、および映像符号化部については、それぞれ独立して動作し、符号化処理を保証し得る専用ハードウェアを用いることによって映像音声の符号化を実現していた。これに対して、専用ハードウェアを用いず、汎用CPUを用いたマルチタスクオペレーティングシステム上で動作するソフトウェアプログラムとして音声、および映像の符号化部を実現することは、きわめて困難である。

【0063】なぜなら、マルチタスクオペレーティングシステム上では、各符号化部もそれぞれタスクとして動作するものであり、同様に他のソフトウェア（通信処理等を行う常駐プログラムなど）もタスクとして機能している場合には、当該他のタスクがある期間だけCPUの動作時間を奪ってしまうことが起こる。その期間には符号化処理は停止するため、常に、符号化ソフトウェアが、音声または映像を十分に処理し得るという保証をすることができない。従って、音声や映像のとぎれ等の再生トラブルを生じない、良好な符号化結果が常に得られるとは限らないこととなる。

【0064】また、映像音声処理対象とする場合には、他のタスク以外にも問題は存在する。すなわち、マルチタスクオペレーティングシステム上で、映像符号化と音声符号化とは別のタスクとして処理されることが常であるため、互いに上記の他のタスクとして影響を与えあうことになるためである。例えば、映像は一樣なものであることは期待できず、入力される映像情報を構成する静止画は時々刻々変化するものとなる。一連の映像のある部分については、きわめて符号化（圧縮）しやすいものであるため、その部分の処理については符号化処理に時間のかかってしまうという可能性もある。この場合、オペレーティングシステム上では、他のソフトウェアがまったく動作していなかったとしても、映像符号化

部に大量のCPU時間をとられ、音声符号化部の処理が遅滞することによって、音声の途切がある符号化結果しか得られないという事態に陥ることがあり得る。

【0065】もう1つの問題は、汎用計算機上のソフトウェア実行により、映像音声符号化装置を構成しようとした場合、AやBの場合と同様に、当該ソフトウェアは、様々なハードウェア能力のコンピュータシステム上で実行される可能性があることに起因する。そのため、前出の問題点、すなわち平均的にみれば音声も映像も符号化できる能力がある場合において、ある期間には符号化に割り当てられる計算機能力が少なくなることによって起因する問題とは別に、当該ソフトウェアが実行されるハードウェアにおいて、そもそも、十分な計算機能力を有しない場合に、当該ソフトウェア設計時の当初の設定値のままでは、映像と音声の符号化処理が行えないという事態が発生する可能性がある。このような場合、動作するコンピュータシステムにあわせ、すみやかに映像符号化が消費する計算機能力を低減しなければ、良好な符号化結果が得られず、再生時に音声途切れが生じてしまうという事態を招く。

【0066】もちろん、音声符号化処理の影響により、映像符号化の符号化結果に不具合が生じる可能性もあるが、一般に同時間に相当する映像と音声とを比較すると、映像の方がデータ量が多いものであり、また映像データの欠落よりも音声データの欠落の影響が、再生時の影響が大きいことから、音声符号化における問題点の方が、比重が大きいものと言え、音声途切れの防止に対する要請がより大きいと一般に言い得るものである。

【0067】以上、A～Cより、パーソナルコンピュータ等の汎用計算機上で符号化ソフトウェアを実行することにより、映像、音声、または映像音声を、とりこみに伴ってのリアルタイム符号化処理をしようとする場合には、以下の問題点があるものと言える。

【0068】（1）当該ソフトウェアを実行するハードウェアの性能の影響が大きい。ハードウェア性能が低いならば、良好な符号化結果が得られなくなり、ハードウェア性能が高いならば、装置資源を活用できなくなる可能性がある。

（2）当該ソフトウェアが、マルチタスクオペレーティングシステム上で実行される場合、他のタスクの影響が大きい。他のタスクによる装置資源の占有の大小が、

（1）におけるハードウェア性能の高低と実質的に同様の影響を与えることとなる。

（3）さらに、映像音声処理対象とする場合には、映像符号化と音声符号化とが互いに他のタスクとして影響を与えあうということが起こり得る。

【0069】本発明は、かかる事情に鑑みてなされたものであり、映像取り込みにともなうリアルタイムでの映像符号化処理を行う符号化方法において、当該符号化方法を実施する計算機の基本的な能力に対応して、解

像度や符号化タイプを含む符号化パラメータを適切に設定して、装置資源を活用して、良好な符号化結果を得ることの可能な映像符号化方法を提供することを目的とする。

【0070】また、本発明は、映像取り込みにともなったリアルタイムでの映像符号化処理を行う符号化方法において、当該符号化方法を実施する計算機のその時点での能力に対応して、解像度や符号化タイプを含む符号化パラメータを適切に設定して、装置資源を活用して、良好な符号化結果を得ることの可能な映像符号化方法を提供することを目的とする。

【0071】また、本発明は、音声取り込みにともなったリアルタイムでの音声符号化処理を行う符号化方法において、当該符号化方法を実施する計算機の基本的な能力に対応して、符号化処理の制御を行い、装置資源を活用して、良好な符号化結果を得ることの可能な音声符号化方法を提供することを目的とする。

【0072】また、本発明は、音声取り込みにともなったリアルタイムでの音声符号化処理を行う符号化方法において、当該符号化方法を実施する計算機のその時点での能力に対応して、符号化処理の制御を行い、装置資源を活用して、良好な符号化結果を得ることの可能な音声符号化方法を提供することを目的とする。

【0073】また、本発明は、音声取り込みにともなったリアルタイムでの音声符号化処理を行う符号化方法において、当該符号化方法を実施する計算機の基本的な能力に対応して、心理聴覚分析の代替処理を実行し、装置資源を活用して、良好な符号化結果を得ることの可能な音声符号化方法を提供することを目的とする。

【0074】また、本発明は、映像音声取り込みにともなったリアルタイムでの映像符号化、および音声符号化処理を行う符号化方法において、当該符号化方法を実施する計算機の基本的な能力に対応して、映像符号化処理の制御を行い、装置資源を活用して、音途切れのない良好な符号化結果を得ることの可能な映像音声符号化方法を提供することを目的とする。

【0075】また、本発明は、映像音声取り込みにともなったリアルタイムでの映像符号化、および音声符号化処理を行う符号化方法において、当該符号化方法を実施する計算機のその時点での能力に対応して、映像符号化処理の制御を行い、装置資源を活用して、音途切れのない良好な符号化結果を得ることの可能な音声符号化方法を提供することを目的とする。

【0076】また、本発明は、上記のような映像符号化方法、音声符号化方法、および映像音声符号化方法を実行する映像符号化装置、音声符号化装置、および映像音声符号化装置を提供することを目的とする。

【0077】また、本発明は、パーソナルコンピュータ等の汎用計算機において実行することで、上記のような映像符号化方法、音声符号化方法、および映像音声符号

化方法を実現できる映像符号化プログラム、音声符号化プログラム、および映像音声符号化プログラムを記録した記録媒体を提供することを目的とする。

【0078】

【課題を解決するための手段】上記目的を達成するため、請求項1にかかる映像符号化方法は、映像を符号化する映像符号化方法において、映像がデジタル化された、複数の静止画像情報からなる原映像情報に対して、上記静止画像情報の1つまたは複数、後述する符号化パラメータに従って符号化する映像符号化ステップと、原映像情報の有する解像度、符号化によって得られる符号化データを再生する際に要求されるフレームレート、上記映像符号化ステップを実行する符号化装置の処理能力を示す処理性能、または上記映像符号化ステップにおける符号化処理の処理量に影響する1つ、もしくは複数の符号化パラメータのうちいずれか1つ以上に基づいて、1つ以上の上記符号化パラメータを決定する符号化パラメータ決定ステップとを実行するものである。

【0079】また、請求項2にかかる映像符号化方法は、請求項1の方法において、当該映像符号化方法の、上記映像符号化ステップを実行する符号化装置の処理能力を判断して、判断結果を出力する処理能力判断ステップをさらに実行するものである。

【0080】また、請求項3にかかる映像符号化方法は、請求項1または2の方法において、上記符号化パラメータは、上記原映像情報に対して行う符号化処理における解像度、フレーム内符号化、もしくは予測符号化を示す符号化タイプ、または上記予測符号化に用いる動きベクトルを検出する際の検出範囲のうち1つ以上を含むものである。

【0081】また、請求項4にかかる映像符号化方法は、請求項2の方法において、上記処理能力判断ステップでは、当該映像符号化方法の有する制御装置の種類に基づいて上記判断を行うものである。

【0082】また、請求項5にかかる映像符号化方法は、請求項2の方法において、上記処理能力判断ステップでは、上記符号化ステップにおける符号化処理の所要時間に基づいて上記判断を行うものである。

【0083】また、請求項6にかかる映像符号化方法は、請求項2の方法において、上記処理能力判断ステップでは、上記入力される原映像情報を一時蓄積し、該蓄積にあたっては、上記原映像情報を構成する一連の静止画像情報を順次保存していくとともに、上記符号化ステップにおいて読み出されて、上記符号化処理が行われた静止画像情報を順次廃棄する映像バッファリングステップと、上記映像バッファリングステップにおける上記一連の静止画像情報の保存を、上記与えられたフレームレートに基づいて決定される一定のフレームレートにおいて行うように制御するフレームレート制御ステップとを実行し、上記映像バッファリングステップにおいて一時

蓄積された上記原映像情報の蓄積量に基づいて上記判断を行うものである。

【0084】また、請求項7にかかる音声符号化方法は、音声に対して、帯域分割符号化方式により符号化を行う音声符号化方法において、符号化処理に用いる数値である、設定周波数 f_s と、変換定数 n とを記憶する記憶ステップと、符号化の対象である音声を入力する音声入力ステップと、上記記憶した設定周波数 f_s に基づいて決定されるサンプリング周波数を用いて、サンプリング音声データを作成する入力音声サンプリングステップと、上記設定周波数 f_s をサンプリング周波数として用いた場合に得られるサンプリング音声データの個数を m 個とし、上記変換定数 n に基づいて定められる数を m' として、 m' 個のサンプリング音声データを含む、 m 個の音声データからなる変換音声データを出力する音声データ変換ステップと、上記変換音声データを、帯域分割して M 個の帯域信号を得る帯域分割ステップと、上記記憶した設定周波数 f_s と変換定数 n とから得られる周波数 $f_s/2n$ を制限周波数として、上記帯域信号のうち、制限周波数以下の帯域信号にのみ符号化ビットを割り当てる符号化ビット割り当てステップと、上記割り当てた符号化ビットに基づいて量子化を行う量子化ステップと、上記量子化したデータを符号化データとして出力する符号化ステップと、上記出力される符号化データを記録する符号化データ記録ステップとを実行するものである。

【0085】また、請求項8にかかる音声符号化方法は、請求項7の方法において、上記入力音声サンプリングステップでは、上記記憶した設定周波数 f_s をサンプリング周波数として、上記入力された音声のサンプリング処理により、 m 個のサンプリング音声データを作成するものであり、上記音声データ変換ステップでは、上記 m 個のサンプリング音声データより、 $(n-1)$ 個おきにサンプリング音声データを抽出し、2つの隣接する上記抽出したサンプリング音声データの間に、 $(n-1)$ 個の音声データを挿入して、 m 個の変換音声データに変換するものである。

【0086】また、請求項9にかかる音声符号化方法は、請求項8の方法において、上記音声データ変換ステップでは、上記抽出したサンプリング音声データがそれぞれ n 個ずつ連続する変換音声データを作成するものである。

【0087】また、請求項10にかかる音声符号化方法は、請求項7の方法において、上記入力音声サンプリングステップでは、上記記憶した設定周波数 f_s と変換定数 n とから得られる周波数 f_s/n をサンプリング周波数として、上記入力された音声のサンプリング処理により、 m/n 個のサンプリング音声データを作成するものであり、上記音声データ変換ステップでは、上記サンプリング音声データに基づき、2つの隣接するサンプリ

ング音声データの間に $(n-1)$ 個の音声データを挿入して、 m 個の変換音声データに変換するものである。

【0088】また、請求項11にかかる音声符号化方法は、請求項10の方法において、上記音声データ変換ステップでは、上記 m/n 個のサンプリング音声データが、それぞれ n 個ずつ連続する変換音声データを作成するものである。

【0089】また、請求項12にかかる音声符号化方法は、請求項7ないし11のいずれかの方法において、上記サンプリング音声データを、入力バッファに一時的に保持する音声バッファリングステップと、上記入力バッファのデータ量を調べて、これを予め設定した値と比較し、上記比較の結果に基づいて、上記レジスタに記憶された上記変換定数 n の値を変更する入力バッファ監視ステップとを実行し、上記入力音声サンプリングステップでは、上記サンプリング音声データを上記入力バッファに書き込むものであり、上記音声データ変換ステップでは、上記入力バッファよりサンプリング音声データを読み出して、これを上記変換するものである。

【0090】また、請求項13にかかる音声符号化方法は、請求項7ないし11のいずれかの方法において、上記符号化ステップにおいて出力される単位時間当たりの符号化データ量を調べて、これを予め設定した値と比較し、上記比較の結果に基づいて、上記レジスタに記憶された上記変換定数 n の値を変更する符号化データ監視ステップを実行するものである。

【0091】また、請求項14にかかる音声符号化方法は、音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化方法において、上記符号化に用いる制御定数を記憶する制御定数記憶ステップと、入力音声をサンプリング処理して、サンプリングデータを出力するサンプリングステップと、上記サンプリングステップで得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割ステップと、上記帯域分割ステップで得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当てステップと、上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化ステップと、上記量子化ステップで得られた量子化値に基づき、符号化データを出力する符号化ステップと、上記記憶した制御定数に基づいて、上記帯域分割ステップ、上記符号化ビット割り当てステップ、上記量子化ステップ、および上記符号化ステップにおけるデータ処理を制御する符号化処理制御ステップとを実行するものである。

【0092】また、請求項15にかかる音声符号化方法は、請求項14の方法において、上記制御定数記憶ステップでは、上記制御定数として、単位期間判定定数 k を単位期間判定定数レジスタに記憶するものであり、上記符号化処理制御ステップは、上記帯域分割ステップでの

1回の帯域分割処理で対象とするサンプリングデータ数を p とし、 p 個のサンプリングデータに相当する時間を単位期間として、上記出力されるサンプリングデータの p 個ごとに、相当する単位期間が符号化対象期間であるか符号化対象外期間であるかの判定を、上記記憶した単位期間判定定数に基づいて行い、上記単位期間が上記符号化対象期間と判定されたときのみ、該単位期間のサンプリングデータが上記帯域分割ステップに出力されるよう制御し、上記単位期間が上記符号化対象外期間と判定されたときは、上記符号化ステップにおいて、予め記憶した固定的符号化データを符号化データとして出力するよう制御する判定制御ステップであるものである。

【0093】また、請求項16にかかる音声符号化方法は、請求項15の方法において、上記判定制御ステップでは、 i 番目の単位期間を t_i として、上記記憶した単位期間判定定数 k と任意の整数 n とから $i = n \times k + 1$ が成立するとき、上記単位期間 t_i が上記符号化対象期間であると判定するものである。

【0094】また、請求項17にかかる音声符号化方法は、請求項14の方法において、上記制御定数記憶ステップでは、上記制御定数として、演算処理判定定数 q を演算処理判定定数レジスタに記憶するものであり、上記符号化処理制御ステップは、上記帯域分割ステップに内包され、上記記憶した演算処理判定定数 q に基づいて、上記帯域分割ステップにおける演算処理を途中で打ち切るように制御する演算処理中止ステップであるものである。

【0095】また、請求項18にかかる音声符号化方法は、請求項17の方法において、上記演算処理中止ステップでは、上記帯域分割ステップにおける基本低域通過フィルタの演算処理を、該フィルタの両端ステップ分については途中で打ち切るように制御するものである。

【0096】また、請求項19にかかる音声符号化方法は、請求項14の方法において、上記制御定数記憶ステップでは、上記制御定数として、帯域選択定数 r を帯域選択定数レジスタに記憶するものであり、上記符号化処理制御ステップは、上記帯域分割ステップが出力する帯域信号データのうち、上記記憶した帯域選択定数 r に基づいて選択したもののみに対して、上記符号化ビット割り当てステップと上記量子化ステップとにおける処理を実行するよう制御する帯域間引きステップであるものである。

【0097】また、請求項20にかかる音声符号化方法は、請求項19の方法において、上記帯域間引きステップでは、上記帯域分割ステップで得られた M 個の帯域信号データ出力から、上記記憶した帯域選択定数である r 個おきに帯域信号データを選択するものである。

【0098】また、請求項21にかかる音声符号化方法は、請求項14ないし20のいずれかの方法において、音声符号化におけるデータ処理の状況を取得し、該取得

した状況に応じて、上記記憶した上記制御定数の値を変更する処理状況監視ステップを実行するものである。

【0099】また、請求項22にかかる音声符号化方法は、請求項21の方法において、上記処理状況監視ステップでは、サンプリングデータを入力バッファに一時蓄積する音声バッファリングステップと、上記入力バッファに保持されるデータの量を予め設定した値と比較し、上記比較の結果に基づいて上記制御定数変更を行う入力監視ステップとを実行するものである。

【0100】また、請求項23にかかる音声符号化方法は、請求項21の方法において、上記処理状況監視ステップは、上記符号化ステップにおいて単位時間あたりに出力される上記符号化データの量を、予め設定した値と比較し、上記比較の結果に基づいて上記制御定数の値を変更する符号化監視ステップであるものである。

【0101】また、請求項24にかかる音声符号化方法は、音声デジタル化された原音声情報に対して、帯域分割符号化方式を用いて符号化を行う音声符号化方法において、入力音声をサンプリング処理して、サンプリングデータを出力するサンプリングステップと、上記サンプリングステップで得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割ステップと、上記帯域分割ステップで得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当てステップと、上記符号化ビット割り当てステップにおける割り当てを心理聴覚分析代替制御方式により制御するビット割り当て制御ステップと、上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化ステップと、上記量子化ステップで得られた量子化値に基づき、符号化データを出力する符号化ステップとを実行するものである。

【0102】また、請求項25にかかる音声符号化方法は、請求項24の方法において、上記ビット割り当て制御ステップは、上記帯域分割ステップで得られた帯域信号データに対して、心理聴覚分析代替制御方式により予め定められたビット割り当て順に従って、符号化ビット割り当てを行うよう制御する順次ビット割り当てステップであるものである。

【0103】また、請求項26にかかる音声符号化方法は、請求項24の方法において、上記ビット割り当て制御ステップは、上記帯域分割ステップで得られた帯域信号データに対して、心理聴覚分析代替制御方式により予め定められた各帯域への重み付けと、各帯域信号データの有する出力レベルとに基づいた符号化ビット割り当てを行うよう制御する帯域出力適応ビット割り当てステップであるものである。

【0104】また、請求項27にかかる音声符号化方法は、請求項24の方法において、上記ビット割り当て制御ステップは、上記帯域分割ステップで得られた帯域信

号データに対して、心理聴覚分析代替制御方式により予め定められた各帯域への重み付けと、各帯域毎のビット割り当て数に対する重み付けと、各帯域信号データの有する出力レベルとに基づいた符号化ビット割り当てを行うよう制御する改良型帯域出力適応ビット割り当てステップであるものである。

【0105】また、請求項28にかかる音声符号化方法は、請求項24の方法において、上記ビット割り当て制御ステップは、上記帯域分割ステップで得られた帯域信号データに対して、帯域信号データごとに最小可聴限界値との比較を行い、上記比較により最小可聴限界未満と判定された帯域信号データにはビット割り当てを行わず、他の帯域に対してのビット割り当てを増加するよう制御する最小可聴限界比較ステップであるものである。

【0106】また、請求項29にかかる映像音声符号化方法は、映像と音声とを符号化するにあたり、上記2つの符号化処理に含まれる処理過程の一部または全部を、共通の計算機資源を用いて実行する映像音声符号化方法において、単位時間毎の静止画像を表す複数の静止画像情報からなる原映像情報と、音声を表す原音声情報とから構成される映像音声情報が入力されたとき、上記原音声情報を一時的に蓄積する音声バッファリングステップと、上記音声バッファリングステップにおいて蓄積された原音声情報を読み出し、この読み出した上記原音声情報を符号化処理し、符号化音声情報を出力する音声符号化ステップと、映像符号化の負荷程度を表す符号化負荷基準情報を用いて、当該映像音声符号化処理についての処理能力を判断し、その判断の結果に基づいて、後述する映像符号化ステップにおける原映像情報に対する符号化を制御する符号化負荷評価ステップと、上記符号化負荷評価ステップにおける制御に従って、入力された上記原映像情報を構成する静止画像情報を符号化処理し、符号化映像情報を出力する映像符号化ステップとを実行するものである。

【0107】また、請求項30にかかる映像音声符号化方法は、請求項29の方法において、上記符号化負荷評価ステップは、上記原映像情報を構成する静止画像情報が入力されたとき、上記音声バッファリングステップにおいて蓄積された原音声情報の総量と、上記符号化負荷基準情報とに基づいて符号化負荷評価情報を求め、上記符号化負荷評価情報を予め設定された負荷限度と比較して、上記符号化負荷評価情報が上記負荷限度に達していない場合に静止画像情報を出力し、上記符号化負荷評価情報が上記負荷限度に達した場合に、上記静止画像情報を破棄するものである。

【0108】また、請求項31にかかる映像音声符号化方法は、請求項29の方法において、アナログ映像情報を入力し、後述する映像解像度情報が出力されたとき、上記アナログ映像情報を複数の離散的デジタル画素情報からなり、上記映像解像度情報に従う解像度を持つ複数

の静止画像情報で構成される原映像情報に変換し、上記映像符号化ステップにおいて処理されるよう出力する映像キャプチャステップを実行するものであり、上記符号化負荷評価ステップでは、上記音声バッファリングステップにおいて蓄積された原音声情報の総量と、映像符号化の負荷程度を表す符号化負荷基準情報とに基づいて符号化負荷評価情報を求め、上記符号化負荷評価情報に基づいて、映像符号化に用いる映像の解像度を表す映像解像度情報を求め、上記映像解像度情報を出力するものであり、上記映像符号化ステップでは、上記映像解像度情報が出力されたとき、上記映像解像度情報に従って上記静止画像情報に対して符号化処理を行い、符号化映像情報を出力するものである。

【0109】また、請求項32にかかる映像音声符号化方法は、請求項29の方法において、上記符号化負荷評価ステップでは、符号化負荷評価情報を上記映像符号化ステップにおいて処理されるよう出力するものであり、上記映像符号化ステップでは、上記静止画像情報に対して、上記出力された符号化負荷評価情報を用いて計算される処理量だけ符号化処理を行い、符号化映像情報として出力するものである。

【0110】また、請求項33にかかる映像音声符号化方法は、請求項29ないし31のいずれかの方法において、上記音声符号化ステップでは、上記音声バッファリングステップにおいて蓄積された原音声情報を読み出し、この読み出した上記原音声情報の総量を計算して処理済み音声情報量として出力し、その後、上記原音声情報を符号化処理して符号化音声情報として出力するものであり、上記符号化負荷評価ステップでは、経過時間と、上記原音声情報の時間当たりの入力量に基づいて原音声入力量を求め、この原音声入力量と上記処理済み音声情報量との差である予測音声バッファ量を求め、上記予測音声バッファ量を用いて、上記符号化負荷評価情報を求めるものである。

【0111】また、請求項34にかかる映像音声符号化方法は、請求項29ないし31のいずれかの方法において、上記符号化負荷評価ステップでは、上記静止画像情報が入力されたとき、経過時間と、上記原音声情報の時間当たりの入力量とに基づいて原音声入力量を求め、かつ、上記音声符号化ステップにおいて出力された符号化音声情報の総量に基づいて処理済み音声情報量を求め、さらに、上記求めた原音声入力量と上記求めた処理済み音声情報量との差である予測音声バッファ量を求めた後、上記予測音声バッファ量を用いて、上記符号化負荷評価情報を求めるものである。

【0112】また、請求項35にかかる映像音声符号化方法は、請求項29ないし31のいずれかの方法において、上記符号化負荷評価ステップにおける、上記判断の結果の変動を監視し、上記変動に対応して、上記符号化負荷基準情報を設定するものである。

【0113】また、請求項36にかかる映像符号化装置は、映像を符号化する映像符号化装置において、映像がデジタル化された、複数の静止画像情報からなる原映像情報に対して、上記静止画像情報の1つまたは複数、後述する符号化パラメータに従って符号化する映像符号化手段と、1つ以上の解像度を一の符号化パラメータとし、フレーム内符号化、順方向予測符号化、逆方向予測符号化、及び双方向予測符号化の各タイプを含む符号化タイプのうち1つ以上の符号化タイプを他の符号化パラメータとして、上記符号化手段の処理量を決定するものである符号化パラメータを、与えられたフレームレートに基づいて決定する符号化パラメータ決定手段とを備えたものである。

【0114】また、請求項37にかかる音声符号化装置は、音声に対して、帯域分割符号化方式により符号化を行う音声符号化装置において、符号化処理に用いる数値である、設定周波数 f_s と、変換定数 n とを記憶するレジスタと、符号化の対象である音声を入力する音声入力手段と、上記記憶した設定周波数 f_s に基づいて決定されるサンプリング周波数を用いて、サンプリング音声データを作成する入力音声サンプリング手段と、上記設定周波数 f_s をサンプリング周波数として用いた場合に得られるサンプリング音声データの個数を m 個とし、上記変換定数 n に基づいて定められる数を m' として、 m' 個のサンプリング音声データを含む、 m 個の音声データからなる変換音声データを出力する音声データ変換手段と、上記変換音声データを、帯域分割して M 個の帯域信号を得る帯域分割手段と、上記記憶した設定周波数 f_s と変換定数 n とから得られる周波数 $f_s/2n$ を制限周波数として、上記帯域信号のうち、制限周波数以下の帯域信号にのみ符号化ビットを割り当てる符号化ビット割り当て手段と、上記割り当てた符号化ビットに基づいて量子化を行う量子化手段と、上記量子化したデータを符号化データとして出力する符号化手段と、上記出力される符号化データを記録する符号化データ記録手段とを備えたものである。

【0115】また、請求項38にかかる音声符号化装置は、音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化装置において、上記符号化に用いる制御定数を記憶する制御定数記憶手段と、入力音声をサンプリング処理して、サンプリングデータを出力するサンプリング手段と、上記サンプリング手段で得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割手段と、上記帯域分割手段で得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当て手段と、上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化手段と、上記量子化手段で得られた量子化値に基づき、符号化データを出力する符号化手段と、上記記憶した制御定数に基づいて、上記

帯域分割手段、上記符号化ビット割り当て手段、上記量子化手段、および上記符号化手段におけるデータ処理を制御する符号化処理制御手段とを備えたものである。

【0116】また、請求項39にかかる音声符号化装置は、音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化装置において、入力音声をサンプリング処理して、サンプリングデータを出力するサンプリング手段と、上記サンプリング手段で得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割手段と、上記帯域分割手段で得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当て手段と、上記符号化ビット割り当て手段における割り当てを心理聴覚分析代替制御方式により制御するビット割り当て制御手段と、上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化手段と、上記量子化手段で得られた量子化値に基づき、符号化データを出力する符号化手段とを備えたものである。

【0117】また、請求項40にかかる映像音声符号化装置は、映像と音声とを符号化するにあたり、上記2つの符号化処理に含まれる処理過程の一部または全部を、共通の計算機資源を用いて実行する映像音声符号化装置において、単位時間毎の静止画像を表す複数の静止画像情報からなる原映像情報と、音声を表す原音声情報とから構成される映像音声情報が入力されたとき、上記原音声情報を一時的に蓄積する音声バッファリング手段と、上記音声バッファリング手段において蓄積された原音声情報を読み出し、この読み出した上記原音声情報を符号化処理し、符号化音声情報を出力する音声符号化手段と、映像符号化の負荷程度を表す符号化負荷基準情報を用いて、当該映像音声符号化装置の処理能力を判断し、その判断の結果に基づいて、後述する映像符号化手段に対しての上記原映像情報の出力を制御する符号化負荷評価手段と、上記符号化負荷評価手段の制御に従って、上記原映像情報を構成する静止画像情報が入力されたとき、上記静止画像情報を符号化処理し、符号化映像情報を出力する映像符号化手段とを備えたものである。

【0118】また、請求項41にかかる映像符号化プログラム記録媒体は、映像を符号化処理する映像符号化プログラムを記録した記録媒体において、映像がデジタル化された、複数の静止画像情報からなる原映像情報に対して、上記静止画像情報の1つまたは複数、後述する符号化パラメータに従って符号化する映像符号化ステップと、1つ以上の解像度を一の符号化パラメータとし、フレーム内符号化、順方向予測符号化、逆方向予測符号化、及び双方向予測符号化の各タイプを含む符号化タイプのうち1つ以上の符号化タイプを他の符号化パラメータとして、上記符号化ステップの処理量を決定するものである符号化パラメータを、与えられたフレームレートに基づいて決定する符号化パラメータ決定ステップとを

実行する符号化プログラムを記録したものである。

【0119】また、請求項42にかかる音声符号化プログラム記録媒体は、音声に対して、帯域分割符号化方式により符号化を行う音声符号化プログラムを記録した記録媒体において、符号化処理に用いる数値である、設定周波数 f_s と、変換定数 n とを記憶する記憶ステップと、符号化の対象である音声を入力する音声入力ステップと、上記記憶した設定周波数 f_s に基づいて決定されるサンプリング周波数を用いて、サンプリング音声データを作成する入力音声サンプリングステップと、上記設定周波数 f_s をサンプリング周波数として用いた場合に得られるサンプリング音声データの個数を m 個とし、 $m \geq m'$ である、上記変換定数 n に基づいて定められる数を m' として、 m' 個のサンプリング音声データを含む、 m 個の音声データからなる変換音声データを出力する音声データ変換ステップと、上記変換音声データを、帯域分割して M 個の帯域信号を得る帯域分割ステップと、上記記憶した設定周波数 f_s と変換定数 n とから得られる周波数 $f_s / 2n$ を制限周波数として、上記帯域信号のうち、制限周波数以下の帯域信号にのみ符号化ビットを割り当てる符号化ビット割り当てステップと、上記割り当てた符号化ビットに基づいて量子化を行う量子化ステップと、上記量子化したデータを符号化データとして出力する符号化ステップと、上記出力される符号化データを記録する符号化データ記録ステップとを実行する符号化プログラムを記録したものである。

【0120】また、請求項43にかかる音声符号化プログラム記録媒体は、音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化プログラムを記録した記録媒体において、上記符号化に用いる制御定数を記憶する制御定数記憶ステップと、入力音声をサンプリング処理して、サンプリングデータを出力するサンプリングステップと、上記サンプリングステップで得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割ステップと、上記帯域分割ステップで得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当てステップと、上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化ステップと、上記量子化ステップで得られた量子化値に基づき、符号化データを出力する符号化ステップと、上記記憶した制御定数に基づいて、上記帯域分割ステップ、上記符号化ビット割り当てステップ、上記量子化ステップ、および上記符号化ステップにおけるデータ処理を制御する符号化処理制御ステップとを実行する符号化プログラムを記録したものである。

【0121】また、請求項44にかかる音声符号化プログラム記録媒体は、音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化プログラムを記録した記録媒体において、入力音声をサンプリング処理して、

サンプリングデータを出力するサンプリングステップと、上記サンプリングステップで得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割ステップと、上記帯域分割ステップで得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当てステップと、上記符号化ビット割り当てステップにおける割り当てを心理聴覚分析代替制御方式により制御するビット割り当て制御ステップと、上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化ステップと、上記量子化ステップで得られた量子化値に基づき、符号化データを出力する符号化ステップとを実行する符号化プログラムを記録したものである。

【0122】また、請求項45にかかる映像音声符号化プログラム記録媒体は、映像と音声とを符号化するにあたり、上記2つの符号化処理に含まれる処理過程の一部または全部を、共通の計算機資源を用いて実行する映像音声符号化プログラムを記録した記録媒体において、単位時間毎の静止画像を表す複数の静止画像情報からなる原映像情報と、音声を表す原音声情報とから構成される映像音声情報が入力されたとき、上記原音声情報を一時的に蓄積する音声バッファリングステップと、上記音声バッファリングステップにおいて蓄積された原音声情報を読み出し、この読み出した上記原音声情報を符号化処理し、符号化音声情報を出力する音声符号化ステップと、映像符号化の負荷程度を表す符号化負荷基準情報を用いて、当該映像音声符号化処理についての処理能力を判断し、その判断の結果に基づいて、後述する映像符号化ステップにおける原映像情報に対する符号化を制御する符号化負荷評価ステップと、上記符号化負荷評価ステップにおける制御に従って、入力された上記原映像情報を構成する静止画像情報を符号化処理し、符号化映像情報を出力する映像符号化ステップとを実行する符号化プログラムを記録したものである。

【0123】

【発明の実施の形態】

実施の形態1. 本発明の実施の形態1による映像符号化方法は、複数の符号化パラメータのうちあるパラメータを定め、設定されたフレームレートと、上記定めたパラメータとに基づいて他のパラメータを決定するものである。

【0124】図1は、本発明の実施の形態1による映像符号化装置の構成を示すブロック図である。図示するように、本実施の形態1による映像符号化装置は、符号化手段101と、符号化パラメータ決定手段102とから構成されており、符号化手段101は、DCT処理手段103、量子化手段104、可変長符号化手段105、ビットストリーム生成手段106、逆量子化手段107、逆DCT処理手段108、および予測画像生成手段109を、また、符号化パラメータ決定手段102は解

像度参照テーブル110を内包している。

【0125】符号化手段101は、映像がデジタル化された、一連の静止画像からなる映像データを入力画像データとして入力し、設定された符号化パラメータに従って符号化処理し、符号化データを出力する。入力画像データを構成する個々の静止画像データをフレーム画像と呼ぶ。また、符号化パラメータは、後述する符号化パラメータ決定手段102から与えられるものであり、符号化タイプを示すパラメータと、解像度を示すパラメータとが含まれている。符号化タイプを示すパラメータは、フレーム内符号化処理、または順方向予測符号化処理を示すものであり、符号化手段101は、当該パラメータに従って、フレーム内符号化、または順方向予測符号化を行う。解像度を示すパラメータは後述するDCT処理手段に入力され、当該解像度において符号化処理が行われることとなる。

【0126】符号化手段101の内部においては、入力画像データに対してまずDCT手段103がDCT（離散コサイン変換）処理を行ってDCT変換データを出力し、次に、量子化手段104が、DCT変換データに対して量子化処理を行って量子化データを出力し、次に可変長符号化手段105が量子化データに対して可変長符号化処理を行うことによって、圧縮符号化された可変長符号化データが作成される。可変長符号化データはビットストリーム生成手段106に入力され、ビットストリーム生成手段106から、伝送や記録を行うことのできるビットストリームとして、装置出力である符号化データが出力される。

【0127】逆量子化手段107は、量子化手段104から出力された量子化データに対して、量子化処理の逆処理である逆量子化処理を行って逆量子化データを出力し、次に逆DCT手段108が逆量子化データに対して、DCT処理の逆処理である逆DCT処理を行って、逆DCT変換データを出力し、逆DCT変換データは予測画像生成手段103に入力され、予測画像データとして出力されることとなる。符号化パラメータに従って、予測画像を用いたフレーム間符号化処理が行われる場合には、この予測画像データと入力画像データとの差分データがDCT手段103に入力されることにより、符号化手段101においては順方向予測符号化が行われることとなる。

【0128】また、本実施の形態1による映像符号化装置では、符号化パラメータ決定手段102は、指定されたフレームレートと符号化パターンとから、内包する解像度参照テーブル110を用いて解像度を決定し、当該決定した解像度を示すパラメータを含む、上記符号化パラメータを符号化手段101に出力する。

【0129】なお、本実施の形態1による映像符号化装置は、パーソナルコンピュータ（PC）において処理制御装置（CPU）の制御により映像符号化プログラムが

実行されることによって実現されるものとし、符号化処理の実行においては、以下の5つの条件が成立するものとする。

【0130】（1）符号化処理時間は、フレーム内符号化処理、順方向予測符号化処理ともに、処理するフレーム画像の解像度に比例した時間を要するものとする。

（2）順方向予測符号化を実行した場合の処理時間は、フレーム内符号化を実行した場合の6倍の時間がかかるものとする。

（3）フレーム内符号化を実行した場合、得られる符号化データ量は入力画像データの1/10の量となり、順方向予測符号化を実行した場合、得られる符号化データのデータ量は入力画像データの1/60の量となるものとする。

（4）本装置を実現するPCのCPUが動作周波数100MHzで動作する場合に、320×240の解像度のフレーム画像をフレーム内符号化を用いて符号化処理した場合には、1/24秒で処理できるものとする。

（5）本装置の処理能力は、本装置に搭載されるCPUの動作周波数に比例するものとする。すなわち、本装置における符号化処理の処理時間は、動作周波数の逆数に比例するものとする。

ここで、本装置に搭載されるCPUの動作周波数は100MHzであり、符号化開始時に指定されるフレームレートは24フレーム/秒、符号化タイプの組み合わせとしての符号化パターンは、すべて「I」のみとするパターンであるパターン1「I」と、2フレームごとに

「I」「P」を繰り返すパターン2「IP」との2種類があるとする。ただし、フレーム内符号化を「I」、順方向予測符号化を「P」で表すものとする。

【0131】以上のような設定のもとに、上述のように構成された本実施の形態1による映像符号化装置の動作を以下に説明する。まず、符号化対象である映像はデジタル化され、一連のフレーム画像として当該符号化装置の符号化手段101に入力される。図2は、符号化手段101の動作を示すフローチャート図である。符号化手段101の動作を、以下に、図2に従って説明する。なお、符号化パラメータ決定手段102は、符号化開始時の最初のフレーム画像に対しては、符号化手段101に対して必ずフレーム内符号化を指示するものとする。

【0132】ステップA01では、符号化パラメータ決定手段102より入力された符号化パラメータについて判断がなされ、フレーム内符号化が指示されていた場合にはステップA02以降の処理が実行され、順方向予測符号化が指示されていた場合には、ステップA07以降の処理が実行される。

【0133】ステップA02以降が実行される場合は、次のようになる。ステップA02でDCT処理手段103は、符号化パラメータ決定手段102が指示する解像度に基づき、入力されたフレーム画像を8画素×8画素

のブロックに分割し、分割したブロックごとに2次元離散コサイン変換して、DCT変換データを出力する。次いで、ステップA03では、量子化手段104は、DCT変換データに対して、ある定められた値を用いて量子化処理を行い、量子化データを出力する。そして、ステップA04で、可変長符号化手段105は、量子化データを可変長符号化し、可変長符号化データを出力する。ステップA05において、ビットストリーム生成手段106は、可変長符号化手段105が出力した可変長符号化データと、符号化パラメータ決定手段102より出力された解像度、および符号化タイプとを用いて、当該映像符号化装置の装置出力として、符号化結果であるビットストリームを生成して出力する。

【0134】ステップA06では、符号化が終了しているか否かが判断され、符号化が終了したと判断されたならば処理は終了する。一方、符号化終了でなければ上記のステップA01に戻り、ステップA01の判断以降が実行される。

【0135】これに対して、ステップA01の判断により、ステップA07以降が実行される場合は次のようになる。まず、ステップA07で逆量子化手段107は、量子化手段104が直前のフレーム画像に対してすでに出力している量子化データを逆量子化し、逆量子化データを出力する。次いでステップA08では、逆DCT処理手段108が、逆量子化データに対して、DCT処理手段103が分割した8画素×8画素のブロックごとに、2次元離散コサイン変換の逆処理である2次元逆離散コサイン変換を実行し、逆DCT変換データを出力する。ステップA09において、予測画像生成手段109は、逆DCT変換データに基づいて予測画像を生成し出力する。

【0136】ステップA10でDCT処理手段103は、入力されたフレーム画像と予測画像生成手段109が出力した予測画像とを、それぞれ指示された解像度に基づき、8画素×8画素のブロックに分割し、分割したブロックごとに、入力されたフレーム画像のデータから予測画像のデータを差し引くことにより差分データを得る。そして、この差分データに対して、分割したブロックごとに2次元離散コサイン変換して、DCT変換データを出力する。DCT変換データが出力された後のステップA11～A14はステップA03からA06と同様に実行される。

【0137】このように、符号化手段101では、入力されたフレーム画像ごとに、ステップA01の判定により、ステップA02～A06か、ステップA07～A14かの処理が行われることとなる。ステップA02～A06はフレーム内符号化であり、ステップA07～A14は直前のフレーム画像に対しての符号化結果を用いた予測画像に基づく順方向符号化処理が行われるものであり、この切り替えはステップA01の判定において、入

力された符号化パラメータに従ってなされるものである。

【0138】(表1)は符号化パラメータ決定手段102が内包する解像度参照テーブル110を示す表である。また、図3は符号化パラメータ決定手段102の動作を示すフローチャート図である。以下に、符号化パラメータを決定して、符号化手段101に出力する符号化パラメータ決定手段102の動作を、表1を参照し、図3のフローに従って説明する。

【0139】

【表1】

入力		出力
フレームレート	符号化パターン	解像度
24	11	320×240
24	1P	160×120

【0140】(表1)に示す解像度参照テーブル110は、符号化に先立ち予め作成しておかれるものである。テーブル作成は、後述する条件を考慮した上で、例えば経験的知識に基づいて、あるいは実験符号化やシミュレーション等の結果を用いて、することができる。

(表1)の「入力」の欄は設定されたフレームレートと指示されるパラメータとを、また「出力」の欄は入力に対応して決定されるパラメータを示している。同表に示すように、本実施の形態1ではフレームレートと符号化パターンとに対応して、解像度が決定される。フレームレートについては固定的に「24(フレーム/秒)」とされているものであり、一方、符号化パターンについては、「11」のパターン1と、「1P」のパターン2とを指示することができる。これらのうち、パターン1「11」はすべてのフレーム画像に対してフレーム内符号化(I)をすることを意味し、パターン2「1P」は、2フレームごとにフレーム内符号化(I)と順方向予測符号化(P)とを繰り返すことを意味する。

【0141】参照テーブルの作成は、次の条件を考慮して行われる。第一に、順方向予測符号化処理は、逆量子化手段107、逆DCT処理手段108、および予測画像生成手段109による処理が付加される分、フレーム内符号化処理よりも処理量が多くなること、第二に、入力画像に対して高解像度での符号化を行った場合は、低解像度で符号化した場合と比べて処理量が多くなることである。これらの条件を考慮し、指定されたフレームレートを実現しつつ、できるだけ高解像度で符号化処理を実行できるように、解像度参照テーブル110は作成されるものである。

【0142】まず、図3のフローのステップB01において、符号化パラメータ決定手段102は、指定されたフレームレートである24フレーム/秒と、符号化パターン(11もしくは1P)とから、解像度参照テーブル110を参照して、符号化を実行するフレーム画像の解像度を決定する。

【0143】次いでステップB02では、符号化パラメータ決定手段102は、符号化手段101に対して、ステップB01で決定した解像度を指示するとともに、指定された符号化パターンを実現できるように、処理対象であるフレーム画像に用いるべき符号化タイプ(1もしくはP)を指示する。

【0144】その後、ステップB03では符号化が終了したか否かが判定され、符号化が終了したと判定されたならば処理は終了する。一方、終了でなければ、ステッ*

入力		出力	符号化結果
フレームレート	符号化パターン	解像度	フレームレート
24	11	320×240	24
24	1P	160×120	27.4

【0147】(表2)は、指示される2つの符号化条件に対して、本実施の形態1の符号化装置において決定される解像度(決定されるパラメータ)と、それらのパラメータを用いた符号化処理の結果として得られたフレームレート(符号化結果)とを示している。(表2)に示す符号化結果の数値については、符号化パターン「11」において、解像度を320×240とした場合に24フレーム/秒で処理できることに基づいて、その他の場合のフレームレートが算出されている。符号化パターンが1Pで解像度が160×120の場合のフレームレートは、Pの処理に1の処理の6倍の時間を要することと、解像度が1/4の場合は1/4の時間で処理できることから、2枚のフレーム画像を符号化するのに(1/24+6/24)÷4=0.073秒を要することが算出でき、これから、27.428フレーム/秒と算出できる。

【0148】比較のため、(表3)に従来の技術による映像符号化装置を用いて符号化を行なった場合の動作結果を示す。

【0149】

【表3】

符号化条件		符号化結果
符号化パターン	解像度	フレームレート
11	640×480	6
11	320×240	24
11	160×120	96
1P	640×480	1.7
1P	320×240	6.9
1P	160×120	27.4

【0150】(表3)においても(表2)の場合と同様の算出がなされており、符号化パターン11において、解像度が320×240の場合に24フレーム/秒で処理できることに基づいて、その他の場合のフレームレートが算出されている。

【0151】従来の技術による映像符号化装置では、符号化結果として得られるフレームレートを考慮せずに、符号化タイプ(パターン)、あるいは解像度を決定していたものである。従って、符号化処理の結果として得ら

* プB02に戻ることによって、符号化手段101に対する符号化パラメータ出力が繰り返される。

【0145】符号化手段101と、符号化パラメータ決定手段102との以上のような動作によって、符号化が実行されるが、(表2)は、本実施の形態1による映像符号化装置において符号化を行なった結果を示す表である。

【0146】

【表2】

れるフレームレートが要望される値に近くなるように設定することが困難であり、不必要な数値となってしまう設定を選定せざるを得ない場合などがあつた。これに比べ、本実施の形態1の映像符号化装置においては、符号化結果であるフレームレートを考慮して、指定された符号化タイプ(パターン)に応じて解像度を決定することで、表2と表3との対比において示されるように、指定されたフレームレートに近いフレームレートを実現しつつ、より高解像度での符号化が実行されていることがわかる。

【0152】このように、本実施の形態1による映像符号化装置によれば、符号化手段101と、解像度参照テーブル110を内包した符号化パラメータ決定手段102とを備えたことで、符号化パラメータ決定手段102は、指定されたフレームレートと符号化タイプとに対応して解像度を決定して、符号化パラメータを符号化手段101に出力し、符号化手段101はこの符号化パラメータに応じて符号化の処理を行うので、要求される条件を実現しつつ、より高解像度での符号化を行うことが可能となる。

【0153】なお、本実施の形態1による映像符号化装置では、指定された符号化パターンに対応して解像度を決定するものとしたが、同様の参照テーブルを用いることによって、指定された解像度に対応して符号化パターン(タイプ)を決定することも可能であり、要求されるフレームレートと符号化パターンとの下で、より圧縮率の高い符号化結果の得られる処理を行うことが可能となる。

【0154】実施の形態2. 本発明の実施の形態2による映像符号化方法は、当該符号化装置の処理能力に対応して、設定されたフレームレートに基づいて符号化パラメータを決定するものであり、制御装置(CPU)の動作周波数により、処理能力を判断するものである。

【0155】図4は、本発明の実施の形態2による映像符号化装置の構成を示すブロック図である。図示するように、本実施の形態2による映像符号化装置は、符号化手段201と、符号化パラメータ決定手段202と、処

理能力判断手段211とから構成されており、符号化手段201は、DCT処理手段203、量子化手段204、可変長符号化手段205、ビットストリーム生成手段206、逆量子化手段207、逆DCT処理手段208、および予測画像生成手段209を、また、符号化パラメータ決定手段202は符号化パターン参照テーブル210を内包している。

【0156】符号化手段201は、実施の形態1による映像符号化装置の符号化手段101と同様であり、符号化パラメータ決定手段102から入力される符号化パラメータに対応して、入力されたフレーム画像に対して、指示された解像度で、かつ、フレーム内符号化(I)、または順方向予測符号化(P)といった指示された符号化タイプで符号化を行う。

【0157】符号化パラメータ決定手段202は、処理能力判断手段211の判断結果に応じて符号化パラメータを決定し、符号化手段201に出力する。処理能力判断手段211は、当該符号化装置の符号化処理能力を判断し、判断結果を符号化パラメータ決定手段202に出力する。本実施の形態2では、処理能力判断手段211は、当該符号化装置の処理能力を示す「CPUの動作周波数」を判断結果として出力するものであり、符号化パラメータ決定手段202は、指定されたフレームレート、および解像度と、判断結果である動作周波数とから符号化パターンを決定し、当該符号化パターンに応じて符号化手段201に符号化タイプを指示するものである。符号化パターンを決定するため、符号化パラメータ決定手段202は、内包する符号化パターン参照テーブル210を用いる。

【0158】なお、本実施の形態2による映像符号化装置においても、実施の形態1と同様に、PCにおける符号化プログラムの実行によって実現されるものとし、実施の形態1に示した条件(1)～(5)が成立するものとする。また、本装置に搭載されるCPUの動作周波数は、100MHz、または166MHzのいずれかであるとし、符号化開始時に指定されるフレームレートは24フレーム/秒であり、入力画像におけるフレーム画像の解像度として、320×240あるいは160×120が指定されるものとする。

【0159】以上のような設定のもとに、上述のように構成された本実施の形態2による映像符号化装置の動作を以下に説明する。入力画像データがフレーム画像ごとに入力され、符号化手段201はこれを符号化処理する。符号化手段201の動作は、実施の形態1において示した符号化手段101と同様である。

【0160】一方、処理能力判断手段211は、本装置の処理能力を判断するために、本装置を実現しているPCに搭載されているCPUの動作周波数を検出し、これを判断結果として符号化パラメータ決定手段202に通知する。符号化パラメータ決定手段202には、動作周

波数100MHz、または166MHzのいずれかを示す判断結果が入力され、符号化パラメータ決定手段202は、この判断結果を用いて符号化パラメータを決定する。

【0161】(表4)は符号化パラメータ決定手段202が内包する符号化パターン参照テーブル210を示す表である。また、図5は符号化パラメータ決定手段202の動作を示すフローチャート図である。以下に、符号化パラメータ決定手段202の動作を、表4を参照し、図5のフローに従って説明する。

【0162】

【表4】

入力			出力
フレームレート	動作周波数	解像度	符号化パターン
24	166	320×240	IIIIIP
24	166	160×120	IPPPPP
24	100	320×240	IIIIIP
24	100	160×120	IPIPIP

【0163】実施の形態1における符号化パラメータ決定手段101の内包する解像度参照テーブル110と同様に、(表4)における符号化パターン参照テーブル210は、後述する条件を考慮して符号化に先立って予め作成されるものである。

【0164】(表4)に示す「入力」欄と「出力」欄との関係については(表1)と同様であり、設定されたフレームレートと解像度、および入力された判断結果である動作周波数の3つに対応して、符号化パターンを決定するものである。ここで、符号化パターンについては、「IIIIIP」はすべてのフレーム画像に対してフレーム内符号化(I)をすることを意味し、「IPIPPIP」は、2フレームごとにフレーム内符号化(I)と順方向予測符号化(P)とを繰り返すことを、「IIIIIP」は6フレームごとにフレーム内符号化(I)を5回繰り返したのち順方向予測符号化(P)を1回実施するという処理を繰り返すことを、また、「IPPPPP」は6フレームごとにフレーム内符号化(I)を1回実施したのち順方向予測符号化(P)を5回繰り返すという処理を繰り返すことを意味する。

【0165】実施の形態1における解像度参照テーブルの設定の場合と同様に、参照テーブルの作成は、次の条件を考慮して行われる。第一に、順方向予測符号化処理は、フレーム内符号化処理と比較して、処理量が多いが高圧縮率で符号化できること、第二に、符号化に際してフレーム画像を高解像度で符号化した場合は、低解像度で符号化した場合と比べて処理量が多くなること、第三に、CPUの動作周波数が高い程その処理能力は高く、符号化処理を短時間で実行できることがその条件である。これらの条件を考慮し、指定されたフレームレートを実現しつつ、できるだけ高い圧縮率で符号化処理を実行できるように、符号化パターン参照テーブル110は

作成される。

【0166】処理能力判断手段211から判断結果が入力されると、符号化パラメータ決定手段202は図5のフローに従って動作する。まず、ステップC01で符号化パラメータ決定手段202は、指定されたフレームレート(24フレーム/秒)と、解像度(320×240、または160×120)と、処理能力判断手段211により入力されるCPUの動作周波数(100MHzもしくは166MHz)とから、符号化パターン参照テーブル210を参照して、符号化を実行する時の符号化パター

10 ーンを決定する。
【0167】続いてステップC02が実行され、符号化パラメータ決定手段202は、符号化手段201に対して符号化パラメータを出力する。符号化パラメータ決定手段202は、ステップC01で決定した符号化パター*

符号化 条件		処理能力 判断 動作 周波数	決定 パラメータ 符号化 パターン	符号化 結果	
フレーム レート	解像度			フレーム レート	符号化 データ量
24	320×240	166	I I I I I P	21.7	0.086
24	160×120	166	I P P P P P	30.8	0.031
24	320×240	100	I I I I I I	24.0	0.100
24	160×120	100	I P I P I P	27.4	0.058

【0170】(表5)においては、指定される符号化条件、および処理能力についての判断結果と、以上から決定される決定パラメータである符号化パターンについては(表4)と同様であり、それぞれの場合について、符号化処理を行った結果(符号化結果)として得られるフレームレート、および符号化データ量を示している。ここで、符号化データ量については、入力画像における1枚のフレーム画像のデータ量を1としたときの、符号化データにおける1枚のフレーム画像に相当する符号化データのデータ量を示している。すなわち、符号化データ量が少ないほど圧縮率が高い。

【0171】なお、(表5)においても、実施の形態1の場合の(表2)と同様に、以下のようにして符号化結果のフレームレートを算出している。すなわち、CPUの動作周波数が100MHz、符号化パターン「I I I I I I」、解像度が320×240の場合に24フレーム/秒で処理できることに基づいて、その他の場合のフレームレートが算出されるものである。例えば、CPUが166MHz、符号化パターン「I I I I I P」、解像度が320×240の場合のフレームレートは、Pの処理

* Nを実現できるように符号化タイプ(IもしくはP)を指示するとともに、指定された解像度を指示する。その後、ステップC03において符号化が終了したか否かが判定され、符号化が終了したと判定されたならば処理は終了する。一方、終了でなければ、ステップC02に戻ることによって、符号化手段201に対する符号化パラメータ出力が繰り返される。

【0168】符号化手段201、符号化パラメータ決定手段202、および処理能力判断手段211の以上のような動作によって、符号化が実行されるが、(表5)は、本実施の形態2による映像符号化装置において符号化を行なった結果を示す表である。

【0169】

【表5】

にIの処理の6倍の時間を要すること、CPUの動作周波数166MHzの場合は100MHzの100/166の時間で符号化を処理できることから、6枚のフレーム画像を符号化するのに $(5/24 + 6/24) \times (100/166) = 0.276$ 秒を要することが算出でき、これから、21.731フレーム/秒と算出できる。同様に、符号化結果として示す1枚のフレーム画像における符号化のデータ量についても、Iで符号化した場合に1/10となること、Pで符号化した場合に1/60となることにそれぞれ基づき算出されている。例えば、パターン「I I I I I P」で符号化した場合には、6枚のフレーム画像の符号化データ量は $(5/10 + 1/60) = 0.517$ になることから、1枚のフレーム画像に対する符号化データ量は、0.086となる。

【0172】比較のため、(表6)に従来の技術による映像符号化装置を用いて符号化を行なった場合の符号化結果を示す。

【0173】

【表6】

符号化条件		処理能力 判断	符号化結果	
解像度	符号化 パターン	動作 周波数	フレーム レート	符号化 データ量
320×240	IIIIII	166	39.8	0.100
320×240	IIIIIP	166	21.7	0.086
320×240	IPPIP	166	11.4	0.058
320×240	IPPPP	166	7.1	0.031
160×120	IIIIII	166	159.4	0.100
160×120	IIIIIP	166	86.9	0.086
160×120	IPPIP	166	45.5	0.058
160×120	IPPPP	166	30.9	0.031
320×240	IIIIII	100	24.0	0.100
320×240	IIIIIP	100	13.1	0.086
320×240	IPPIP	100	6.9	0.058
320×240	IPPPP	100	4.6	0.031
160×120	IIIIII	100	96	0.100
160×120	IIIIIP	100	52.4	0.086
160×120	IPPIP	100	27.4	0.058
160×120	IPPPP	100	18.6	0.031

【0174】なお、(表6)においても、CPUの動作周波数が100MHz、符号化パターン「IIIIII」で、解像度が320×240の場合に24フレーム/秒で処理できることも基づいて、その他の場合の符号化結果であるフレームレートを算出しているものである。また、符号化データ量に関しても、1枚のフレーム画像に対して、Iで符号化した場合に1/10となること、Pで符号化した場合に1/60となることに基づき算出されている。

【0175】従来の技術による映像符号化装置では、符号化結果として得られるフレームレートや、当該符号化装置を構成するハードウェア能力の変動を考慮せずに、符号化タイプ(パターン)、あるいは解像度を決定していたものである。従って、かかる設定に従った符号化処理の結果として得られるフレームレートは、ときには不必要な数値になってしまうなど不具合があった。これに比べ、本実施の形態2の映像符号化装置においては、当該符号化装置の処理能力や、符号化結果であるフレームレートを考慮して、指定された解像度に応じて符号化タイプ(パターン)を決定することで、表5と表6との対比において示されるように、指定されたフレームレートに近いフレームレートを実現でき、かつ、より高圧縮率での符号化が実行されていることがわかる。

【0176】特に、当該符号化装置のハードウェア的能力の変動に対応し得るという利点は、映像符号化プログラムをコンピュータ等において実行することによって、当該映像符号化装置を実現する場合には有用である。

【0177】このように、本実施の形態2による映像符号化装置によれば、符号化手段201と、符号化パターン参照テーブル210を内包した符号化パラメータ決定手段202と、処理能力判断手段211とを備えたことで、符号化パラメータ決定手段202は、指定されたフレームレートと解像度、そして処理能力判断手段211

の出力する判断結果とに対応して符号化パターンを決定して、符号化パラメータを符号化手段201に出力し、符号化手段201はこの符号化パラメータに応じて符号化の処理を行うので、要求される条件を実現しつつ、より高圧縮率の得られる符号化を行うことが可能となる。

【0178】なお、本実施の形態2による映像符号化装置では、指定された解像度に対応して符号化パターンを決定するものとしたが、同様の参照テーブルを用いることによって、指定された符号化パターン(タイプ)に対応して解像度を決定することも可能であり、要求されるフレームレートと符号化パターンとの下で、より高解像度での符号化処理をすることが可能となる。

【0179】また、本実施の形態2では、処理能力の判断については、CPUの動作周波数に基づいて行うものとしているが、CPUあるいはDSP等のプロセッサの品番、バージョン、製造メーカーなどの、装置能力を示す諸要素に基づいて判断することとしてもよく、種々の応用が可能である。

【0180】実施の形態3. 本発明の実施の形態3による映像符号化方法は、当該符号化装置の処理能力に対応して、設定されたフレームレートに基づいて符号化パラメータを決定するものであり、所要処理時間に基づいて処理能力を判断するものである。

【0181】図6は、本発明の実施の形態3による映像符号化装置の構成を示すブロック図である。図示するように、本実施の形態3による映像符号化装置は、符号化手段301と、符号化パラメータ決定手段302と、処理能力判断手段311とから構成されており、符号化手段301は、DCT処理手段303、量子化手段304、可変長符号化手段305、ビットストリーム生成手段306、逆量子化手段307、逆DCT処理手段308、および予測画像生成手段309を、また、符号化パラメータ決定手段302は符号化パターン決定手段31

0を内包している。符号化手段301は、実施の形態1による映像符号化装置の符号化手段101と同様であり、符号化パラメータ決定手段302から入力される符号化パラメータに対応して、入力されたフレーム画像に対して、指示された解像度で、かつ、フレーム内符号化(I)、または順方向予測符号化(P)といった指示された符号化タイプで符号化を行う。

【0182】符号化パラメータ決定手段302は、処理能力判断手段311の判断結果に応じて、内包する符号化パターン決定手段310を用いて符号化パラメータを決定し、符号化手段301に出力する。処理能力判断手段311は、当該符号化装置の符号化処理能力を判断し、判断結果を符号化パラメータ決定手段302に出力する。本実施の形態3では、処理能力判断手段311は、当該符号化装置における符号化処理の平均フレームレートを判断結果として出力するものであり、符号化パラメータ決定手段302は、指定されたフレームレートと、解像度と、判断結果である平均フレームレートとから符号化パターンを決定し、当該符号化パターンに応じて符号化手段301に符号化タイプを指示するものである。符号化パターンを決定するため、符号化パラメータ決定手段302は、符号化パターン決定手段310を用いる。

【0183】なお、本実施の形態3による映像符号化装置においても、実施の形態1と同様に、PCにおける符号化プログラムの実行によって実現されるものとし、実施の形態1に示した条件(1)～(5)が成立するものとする。また、本装置に搭載されるCPUの動作周波数は100MHzであるとする。また、符号化開始時に指定されるフレームレートは8フレーム/秒であり、入力画像におけるフレーム画像の解像度として320×240が指定されるとする。

【0184】以上のような設定のもとに、上述のように構成された本実施の形態3による映像符号化装置の動作を、以下に説明する。入力画像データがフレーム画像ごとに入力され、符号化手段301はこれを符号化処理する。符号化手段301の動作は、実施の形態1において示した符号化手段101と同様である。

【0185】一方、処理能力判断手段311は、本装置の処理能力を判断するために、4つのフレーム画像ごとに、それら4つのフレーム画像を含めた、それら以前のすべてのフレーム画像に対して、符号化手段301が処理するのに要した時間を測定し、当該測定した処理所要時間と、フレーム画像の処理数とから、その時点までの符号化処理の平均フレームレートを算出し、符号化パラメータ決定手段302に通知する。なお、処理能力判断手段は平均フレームレートの初期値としては、要求されたフレームレートである8フレーム/秒を通知するものとする。

【0186】符号化パラメータ決定手段302は、上記

通知された平均フレームレートを用いて符号化パラメータを決定する。符号化パラメータ決定手段302に含まれる符号化パターン決定手段310の動作を以下に説明する。符号化パターン決定手段310は、限定されたいくつかの状態をとるものであり、条件に応じて、それらのとり得る状態間を遷移する有限状態マシンとして動作する。図7は、有限状態マシンとしての符号化パターン決定手段310における(a)状態の遷移を示す状態遷移図、および(b)状態遷移条件を示す図である。符号化パターン決定手段310はS0～S3までの全部で4つの状態をとり、それぞれの状態においては(表7)で示す符号化パターンを出力する。

【0187】

【表7】

状態	符号化パターン
S0	IIII
S1	IIIP
S2	IPIP
S3	IPPP

【0188】(表7)に示す符号化パターンについては、「IIII」は、処理対象である4つのフレーム画像のすべてのフレーム画像に対してフレーム内符号化(I)を実行することを、「IIIP」は該4つのフレーム画像のうち、最初の3つに対してはフレーム内符号化(I)を、最後の1つに対しては順方向予測符号化(P)を実行することを、「IPIP」は該4つのフレーム画像のうち、最初と3つめに対してはフレーム内符号化(I)を、2つ目と4つ目に対しては順方向予測符号化(P)を実行することを、「IPPP」は、該4つのフレーム画像のうち、最初の1つに対してはフレーム内符号化(I)を、残りの3つに対しては、順方向予測符号化(P)を実行することを、それぞれ意味する。

【0189】また、符号化パターン決定手段310の状態の遷移については、4つのフレーム画像ごとに状態遷移についての判定を実行するものであり、判定は、当該判定の直前に処理能力判断手段311から通知された平均フレームレートの値と、指定されたフレームレートの値とを用いて、図9(a)に示す条件に従って行なわれる。ただし、本実施の形態3においては、上記の有限状態マシンとしてとる状態の初期値はS1であるものとする。

【0190】以上のことから、本実施の形態3による映像符号化装置では、符号化が開始するとまず、処理能力判断手段311から初期値「8フレーム/秒」が出力され、符号化パターン決定手段310は初期状態S1であることから、表7に示すように符号化パターンとして「IIIP」を出力する。従って、符号化パラメータ決定手段302は、該パターンを実現できるように符号化パラメータを符号化手段301に出力し、フレーム画像3つに対してフレーム内符号化、次の1つに対して順方

向予測符号化が行われるように制御がなされる。

【0191】この後、処理能力判断手段311から得られる平均フレームレートが、指定されたフレームレートより低いものとなったときには、図7(a)に示すようにS1→S0の遷移がなされることにより、表7に示す符号化パターンは「IIII」に変更され、フレーム内符号化ばかりが行われるようになる。

【0192】一方、処理能力判断手段311から得られる平均フレームレートが指定されたフレームレートよりも高いときには、図7(a)に示すS1→S2の遷移がされて、符号化パターンが「IPIP」に変更され、順方向予測符号化の比率が増すこととなる。

【0193】このような制御をすることによって、処理能力判断手段311の出力する平均フレームレートが指定されたフレームレートよりも小さいときは、当該符号化装置の処理負担が重いものと考えられるため、図7

(a)に示すS3→S0方向の遷移によって、符号化処理*

フレーム画像No.	0→3	4→7	8→11	12→15	16→19	20→23	24→27
指定フレームレート	8						
符号化パターン	IIIP	IIIP	IPIP	IPPP	IPPP	IPIP	IIIP
所要時間	0.375	0.375	0.583	0.792	0.792	0.583	0.375
311の出力	8	10.7	10.7	9	7.5	6.9	6.9
結果(フレームレート)	7.23						
結果(データ量)	0.06						

【0196】同表において、指定されるフレームレートは、固定的に「8」と設定されている。符号化パターン、所要時間、および処理能力判断手段311の出力については、0番目から3番目、4番目から7番目…等の4つずつのフレーム画像の処理ごとに、その値を示すものである。符号化パターンは4つのフレーム画像ごとに符号化処理に用いられる符号化パターンを、所要時間は符号化手段301が4つのフレーム画像の符号化処理に要する時間(秒)を、また、処理能力判断手段311の出力は、所要時間を用いて取得した平均フレームレートを示している。そして、28枚のフレーム画像を符号化した結果として、符号化処理における平均フレームレートと符号化データ量とを示している。

【0197】なお、(表8)においては、CPUの動作周波数が100MHz、符号化パターン「IIII」において、解像度が320×240の場合に、24フレーム/秒で処理できることから、4枚のフレーム画像を処理するのに必要な時間と、符号化結果としてのフレームレートとが算出されている。例えば、0番目から3番目までの4枚のフレーム画像を処理するのに必要な時間は、 $1/24 \times 3 + 6/24 \times 1 = 0.375$ 秒となる。また、符号化結果としては、1フレームが15枚とPフレームが13枚生成されるため、1枚のフレーム画像の符

*における、処理負担の小さなフレーム内符号化の比率を高くし、一方、処理能力判断手段311の出力する平均フレームレートが指定されたフレームレートよりも大きいときは、当該符号化装置の処理能力に余力があるものと考えられるため、図7(a)に示すS0→S3方向の遷移によって、符号化処理における、処理負担の大きな順方向予測符号化の比率を高くして、より高圧縮度の符号化結果が得られるように図るものである。

【0194】このように、処理能力として示される符号化処理の状態に対応して、符号化パラメータが変化させながら、符号化が実行されるが、(表8)は、上記のようにして、28枚の連続したフレーム画像に対して符号化を実施した場合における、符号化の結果を示す表である。

【0195】

【表8】

号化に要した時間は、 $(1/24 \times 15 + 6/24 \times 13) \div 28 = 0.138$ 秒となり、平均フレームレートは7.225フレーム/秒と算出できる。また、符号化結果における1枚のフレーム画像の符号化データ量は、Iで符号化した場合に1/10となること、Pで符号化した場合に1/60となることに基づいて算出されている。その結果、符号化結果としては、1フレームが15枚とPフレームが13枚生成されるため、28枚のフレーム画像の符号化データ量は $(15/10 + 13/60)$ になることから、1枚のフレーム画像に対する符号化データ量は、0.061となる。

【0198】(表9)は、本実施の形態3による映像符号化装置における、符号化パターン決定手段310、および符号化パラメータ決定手段302の機能を説明するための表である。同表においては、上記(表8)における処理対象の28枚のフレーム画像のうち、0番目から11番目までの12枚のフレーム画像について1枚ごとについての、符号化パターン決定手段310、および符号化パラメータ決定手段302の出力を含む、本実施の形態3による映像符号化装置の状態を示すものである。

【0199】

【表9】

フレーム画像No.	0	1	2	3	4	5	6	7	8	9	10	11
310の状態	S 1				S 1				S 2			
310の出力	I I P				I I P				I P I P			
302の出力	I	I	I	P	I	I	I	P	I	P	I	P
所要時間	0.375				0.375				0.583			
311の出力	8.0				10.7				10.7			

【0200】同表において、所要時間と、処理能力判断手段311の出力とは、(表8)と同様である。この処理能力判断手段311の出力に対応して、符号化パターン決定手段310は状態遷移をし、状態に応じて符号化

10

パターンを出力する。そして、符号化パラメータ決定手段302は、出力された符号化パターンに対応して、同表に示すように、1枚のフレーム画像ごとに符号化タイプを出力する。

【0201】同表に示すように、符号化パターン決定手段310の状態は、上記のようにまずS1をとるので、このため表7に示すように符号化パターン「I I I P」が出力される。符号化パラメータ決定手段302はこれに対応して、最初の4つのフレーム画像について「I」

20

「I」「I」「P」を出力する。

【0202】8番目から11番目までの4つのフレーム画像に対しては、符号化パターン決定手段310の状態がS1→S2に遷移したことから、符号化パターン「I P I P」が出力されるので、符号化パラメータ決定手段302の出力は、「I」「P」「I」「P」となる。

【0203】以下、表8に示すように、4つのフレーム画像を処理するごとに、処理能力の判断と、それに対応した符号化パターンの選択がなされ、該パターンに対応した符号化タイプにおいて符号化処理が行われる。

【0204】比較のため、(表10)に従来の技術による映像符号化装置を用いて、28枚のフレーム画像を、予め決められた符号化パターンで符号化した場合における符号化結果を示す。

30

【0205】

【表10】

(a) 符号化パターンI I I Iに対する符号化結果

フレーム画像No.	1→28
符号化パターン	I I I I
結果(フレームレート)	24
結果(データ量)	0.1

(b) 符号化パターンI I I Pに対する符号化結果

フレーム画像No.	1→28
符号化パターン	I I I P
結果(フレームレート)	10.7
結果(データ量)	0.08

(c) 符号化パターンI P I Pに対する符号化結果

フレーム画像No.	1→28
符号化パターン	I P I P
結果(フレームレート)	6.9
結果(データ量)	0.06

(d) 符号化パターンI P P Pに対する符号化結果

フレーム画像No.	1→28
符号化パターン	I P P P
結果(フレームレート)	5.1
結果(データ量)	0.04

【0206】符号化パターンとしては、(表7)で示されている4つのパターンを用いており、符号化結果として、符号化処理のフレームレートと、1枚のフレーム画像に対する符号化データ量とを示している。なお、(表10)においても、CPUの動作周波数が100MHz、符号化パターンI I I Iで、解像度が320×240の場合に24フレーム/秒で処理できることから、その他の場合のフレームレートが算出されている。また、符号化データ量に関しても、1枚のフレーム画像に対して、Iで符号化した場合に1/10となること、Pで符号化した場合に1/60となることに基づき算出されている。

【0207】従来の技術による映像符号化装置では、符号化結果として得られるフレームレートや、当該符号化装置を構成するハードウェア能力の変動を考慮せずに、符号化タイプ(パターン)、あるいは解像度を決定していたものである。従って、符号化処理の結果として得られるフレームレートが、要望される値に近くなるよう設定することは困難であり、不必要な数値となってしまう設定を選定せざるを得ない場合などがあった。これに比べ、本実施の形態3の映像符号化装置においては、当該符号化装置の処理能力や符号化結果であるフレームレートの変動を考慮して、指定された解像度に応じて符号化タイプ(パターン)を決定することで、表9と表10と

50

の対比において示されるように、指定されたフレームレートに近いフレームレートを実現でき、かつ、より高圧縮率での符号化が実行されていることがわかる。

【0208】このように、本実施の形態3による映像符号化装置によれば、符号化手段301と、符号化パターン決定手段310を内包した符号化パラメータ決定手段302と、処理能力判断手段311とを備えたことで、符号化パラメータ決定手段302は、指定されたフレームレートと解像度、そして処理能力判断手段311の出力する判断結果とに対応して符号化パターンを決定し、符号化パラメータを符号化手段301に出力し、符号化手段301はこの符号化パラメータに応じて符号化の処理を行うので、要求される条件を実現しつつ、より高圧縮率の得られる符号化を行うことが可能となる。

【0209】なお、本実施の形態3による映像符号化装置では、指定された解像度に対応して符号化パターンを決定するものとしたが、同様の処理をすることによって、指定された符号化パターン（タイプ）に対応して解像度を決定することも可能であり、要求されるフレームレートと符号化パターンとの下で、より高解像度での符号化処理を行うことが可能となる。

【0210】また、実施の形態1、および2による映像符号化装置では、基本的に符号化開始に際して符号化パラメータを決定し、それ以後は該決定された符号化パラメータに従って符号化処理をするものであるが、本実施の形態3による装置では、符号化の処理を行いながら平均レートを算定し、装置の処理能力として取得される符号化の状況に対応して動的に符号化パラメータを変更し得るものである。従って、本実施の形態3による映像符号化装置では、実施の形態1、および2に比較して若干の処理負担は伴うものの、複数の演算処理等を並行して実行する汎用計算機において、映像取り込みに伴った符号化を実行する際などであって、当該計算機装置の状況が変化するような場合でも、その状況の変化に対応して、適切な符号化条件を設定することが可能となるものである。

【0211】もっとも、あまり状況の変化がないなど、当該符号化装置の処理能力が符号化処理の過程において大きく変動しないと見込まれる場合等には、本実施の形態3による映像符号化装置においても、実施の形態1、および2と同様に、符号化開始に際してパラメータを設定し、以後はその条件で符号化を実施するものとして、制御にかかる処理負担を軽減することも可能である。

【0212】実施の形態4. 本発明の実施の形態4による映像符号化方法は、当該符号化装置の処理能力に対応して、設定されたフレームレートに基づいて符号化パラメータを決定するものであり、一時蓄積するデータの量により処理能力を判断するものである。

【0213】図8は、本発明の実施の形態4における映像符号化方法を実行する映像符号化装置の構成を示すブ

ロック図である。図示するように、本実施の形態4による映像符号化装置は、符号化手段401、符号化パラメータ決定手段402、処理能力判断手段411、バッファ手段412、および入力フレームレート制御手段413から構成されている。符号化手段401は、DCT処理手段403、量子化手段404、可変長符号化手段405、ビットストリーム生成手段406、逆量子化手段407、逆DCT処理手段408、および予測画像生成手段409を、また、符号化パラメータ決定手段402は符号化パターン決定手段410を内包している。

【0214】符号化手段401は、実施の形態1による映像符号化装置の符号化手段101と同様であり、符号化パラメータ決定手段402から入力される符号化パラメータに対応して、入力されたフレーム画像に対して、指示された解像度で、かつ、フレーム内符号化（I）、または順方向予測符号化（P）といった指示された符号化タイプで符号化を行う。実施の形態1～3において、符号化手段101～301はいずれも入力画像データを入力するものであったが、本実施の形態4では、符号化手段401は後述するバッファ手段412よりデータを読み出して符号化処理を行うものである。

【0215】符号化パラメータ決定手段402は、処理能力判断手段411の判断結果に応じて、符号化パラメータを決定し、符号化手段401に出力する。処理能力判断手段411は、当該符号化装置の符号化処理能力を判断し、判断結果を符号化パラメータ決定手段402に出力する。本実施の形態4では、処理能力判断手段411は、後述するバッファ手段412に一時蓄積されるバッファ量を判断結果として出力するものである。また、符号化パラメータ決定手段402は、指定されたフレームレートと、解像度と、判断結果であるバッファ量とから符号化パターンを決定し、当該符号化パターンに応じて符号化手段401に符号化タイプを指示するものである。符号化パターンを決定するため、符号化パラメータ決定手段402は、符号化パターン決定手段410を用いる。

【0216】入力フレームレート制御手段413は、当該映像符号化装置の入力である入力画像データを、一連のフレーム画像として、指定されたフレームレートに対応して、後述するバッファ手段412に出力する。バッファ手段412は、入力画像データを一時蓄積するものであって、入力画像データを一連のフレーム画像として順次保存していくとともに上記符号化手段401により読み込まれたフレーム画像を順次廃棄していく。そして、本実施の形態4では、処理能力判断手段411は、符号化開始時にバッファ手段412に保存されていたフレーム画像の枚数と、現在保存されているフレーム画像の枚数との差分を検出してこれを判断結果として、符号化パラメータ決定手段402に出力するものである。

【0217】なお、本実施の形態4による映像符号化装

置においても、実施の形態1と同様に、PCにおける符号化プログラムの実行によって実現されるものとし、実施の形態1に示した条件(1)～(5)が成立するものとする。また、本装置に搭載されるCPUの動作周波数は100MHzであるとする。また、符号化開始時に指定されるフレームレートは8フレーム/秒であり、入力画像におけるフレーム画像の解像度として320×240が指定されるものとする。

【0218】以上のような設定のもとに、上述のように構成された本実施の形態4による映像符号化装置の動作を以下に説明する。本実施の形態4による映像符号化装置の処理対象である映像が、入力画像データとして入力されると、この入力画像データはまず、入力フレームレート制御手段413に輸入される。入力フレームレート制御手段413は、指定されたフレームレートで、入力画像データを一連のフレーム画像としてバッファ手段412に順次入力していく。本実施の形態4においては、入力フレームレート制御手段413は、上記の8フレーム/秒を該指定されたフレームレートとして処理を行う。

【0219】バッファ手段412は、入力フレームレート制御手段413より入力されたフレーム画像を順次保存してゆき、符号化手段401により読み込まれたフレーム画像を順次廃棄する。すなわち、FIFO(先入れ先出し)方式によりデータを一時蓄積するものである。なお、符号化開始時において、符号化手段401は入力フレームレート制御手段413よりも一定の時間遅れて動作を開始する。すなわち、符号化開始時において、バッファ手段412は、ある一定の枚数のフレーム画像を保存している。これは、バッファのアンダーフロー、すなわち一時蓄積するデータの枯渇により処理が円滑に進行しなくなることを防ぐものである。この段階で処理能力判断手段411は、初期状態におけるバッファ手段412に蓄積されたフレーム画像の枚数を検出し、比較処理のために保持する。

【0220】符号化手段401は、バッファ手段412より一時蓄積されたフレーム画像を読み出して、符号化処理を行う。符号化処理に際しての符号化手段401の動作は、実施の形態1における符号化手段101と同様のものとなる。一方、処理能力判断手段411は、符号化手段401が4枚のフレーム画像を処理するごとに、バッファ手段412に保存されているフレーム画像の枚数を検出し、先に検出して保持している、初期状態におけるバッファ手段412のフレーム画像の枚数との差分を取得して、この差分値を符号化パラメータ決定手段402に通知する。この時、保存されているフレーム画像の枚数が、初期状態におけるフレーム画像の枚数よりも多い場合は、差分を正の値として、少ない場合は差分を負の値として通知するものとする。

【0221】符号化パラメータ決定手段402では、上

記の差分が、符号化パターン決定手段410に輸入される。符号化パターン決定手段410は、実施の形態3における符号化パターン決定手段310と同様に、有限状態マシンとして動作する。図9(a)は、有限状態マシンとして動作する符号化パターン決定手段410の状態遷移図であり、同図(b)は、状態遷移条件を示す図である。符号化パターン決定手段はS0～S3までの、全部で4つの状態をとり、それぞれの状態において、(表11)で示す符号化パターンを出力する。

【0222】

【表11】

状態	符号化パターン
S0	IIII
S1	IIIP
S2	IPIP
S3	IPPP

【0223】(表11)に示す、「IIII」、「IIIP」、「IPIP」、および「IPPP」の各パターンは実施の形態3の(表7)で示したものと同等である。また、符号化パターン決定手段410における状態の遷移は、符号化手段401が4枚のフレーム画像を処理するごとに判定がなされるものであって、判定は、処理能力判断手段411から通知された当該判定の直前の差分値に基づいて、図9(b)に示した条件に従って行なわれる。なお、本実施の形態4においては、有限状態マシンとしての状態の初期値はS1であるとする。

【0224】符号化パラメータ決定手段402は、符号化手段401が4枚のフレーム画像を処理するごとに、符号化パターン決定手段410が出力する符号化パターンを取得して、当該取得した符号化パターンを実現できるように、1枚のフレーム画像ごとに符号化手段401に対して符号化タイプを指示する。また、指定された解像度を符号化手段401に対してそのまま指示する。

【0225】以上のことから、本実施の形態4による映像符号化装置では、符号化が開始するとまず、符号化パターン決定手段410は初期状態S1であることから、表12に示すように符号化パターンとして「IIIP」を出力する。従って、符号化パラメータ決定手段402は、該パターンを実現できるように符号化パラメータを符号化手段401に出力し、フレーム画像3つに対してフレーム内符号化、次の1つに対して順方向予測符号化が行われるように制御がなされる。

【0226】この後、処理能力判断手段411から得られる差分値が、負の値となったときには、図9(a)に示すようにS1→S0の遷移がなされることにより、表12に示す符号化パターンは「IIII」に変更され、フレーム内符号化ばかりが行われるようになる。一方、処理能力判断手段411から得られる差分値が正の値になったときには、図9(a)に示すS1→S2の遷移がされ、符号化パターンが「IPIP」に変更され、順方向

予測符号化の比率が増すこととなる。

【0227】このような制御をすることにより、処理能力判断手段411の出力する差分値が正である、すなわち、バッファ手段412に蓄積されたフレーム画像の枚数が、初期の蓄積枚数より多いときは、当該符号化装置の処理負担が重いものと考えられるため、図9(a)に示すS3→S0方向の遷移によって、符号化処理における、処理負担の小さなフレーム内符号化の比率を高くするように図る。一方、処理能力判断手段411の出力する差分値がである、すなわち、バッファ手段412に蓄

積されたフレーム画像の枚数が、初期の蓄積枚数より多いときは、当該符号化装置の処理能力に余力があるもの*

フレーム画像No.	0→3	4→7	8→11	12→15	16→19	20→23	24→27
指定フレームレート	8						
符号化パターン	IIIP	IIIP	IPIP	IPPP	IPPP	IPIP	IIIP
所要時間	0.375	0.375	0.583	0.792	0.792	0.583	0.375
入力枚数	3.0	3.0	4.7	6.3	6.3	4.7	3.0
出力枚数	4.0	4.0	4.0	4.0	4.0	4.0	4.0
411の出力	0.0	-1.0	-2.0	-1.3	+1.0	+3.3	+4.0
結果(フレームレート)	7.23						
結果(データ量)	0.06						

【0230】同表において、指定されるフレームレートは、固定的に「8」と設定されている。符号化パターン、所要時間、入力枚数、出力枚数、および処理能力判断手段411の出力については、0番目から3番目、4番目から7番目…等の4つずつのフレーム画像の処理ごとに、値を示すものである。符号化パターンは、4つのフレーム画像ごとに符号化処理に用いられる符号化パターンを、所要時間は、符号化手段401が4つのフレーム画像の符号化処理に要する時間(秒)を、入力枚数は、バッファ手段412に対してフレーム画像が入力された枚数を、出力枚数は、バッファ手段412からフレーム画像が出力された枚数を示している。そして、28枚のフレーム画像を符号化した結果として、符号化処理における平均フレームレートと符号化データ量とを示している。

【0231】なお、(表12)においては、CPUの動作周波数が100MHz、符号化パターン「IIII」において、解像度が320×240の場合に24フレーム/秒で処理できることに基づいて、4枚のフレーム画像を処理するのに必要な時間および、符号化結果としてのフレームレートが算出されている。

【0232】例えば、0番目から3番目までの4枚のフレーム画像を処理するのに必要な時間は、 $1/24 \times 3 + 6/24 \times 1 = 0.375$ 秒となる。また、符号化結果としては、Iフレームが15枚とPフレームが13枚

*と考えられるため、図9(a)に示すS0→S3方向の遷移によって、符号化処理における、処理負担の大きな順方向予測符号化の比率を高くして、より高圧縮度の符号化結果が得られるように図るものである。

【0228】このように、蓄積枚数として示される符号化処理の状態に対応して、符号化パラメータが変化させながら、符号化が実行されるが、(表12)は、上記のようにして、28枚の連続したフレーム画像に対して符号化を実施した場合における、符号化の結果を示す表である。

【0229】

【表12】

生成されるため、1枚のフレーム画像の符号化に要した時間は、 $(1/24 \times 15 + 6/24 \times 13) \div 28 = 0.138$ 秒となり、平均フレームレートは7.225フレーム/秒と算出できる。また、符号化結果における1枚のフレーム画像の符号化データ量は、Iで符号化した場合に1/10となること、Pで符号化した場合に1/60となることにそれぞれ基づき算出されている。その結果、符号化結果としては、Iフレームが15枚とPフレームが13枚生成されるため、28枚のフレーム画像の符号化データ量は $(15/10 + 13/60) = 1.717$ になることから、1枚のフレーム画像に対する符号化データ量は、0.061となる。

【0233】(表13)は、本実施の形態4による映像符号化装置における、符号化パターン決定手段410、および符号化パラメータ決定手段402の機能を説明するための表である。同表においては、上記(表12)における処理対象の28枚のフレーム画像のうち、0番目から11番目までの12枚のフレーム画像について1枚ごとについての、符号化パターン決定手段410、および符号化パラメータ決定手段402の出力を含む、本実施の形態3による映像符号化装置の状態を示すものである。

【0234】

【表13】

フレーム画像No.	0	1	2	3	4	5	6	7	8	9	10	11
410の状態	S 1				S 1				S 2			
410の出力	I I P				I I P				I P I P			
402の出力	I	I	I	P	I	I	I	P	I	P	I	P
所要時間	0.375				0.375				0.583			
入力枚数	3.0				3.0				4.7			
出力枚数	4.0				4.0				4.0			
411の出力	0.0				-1.0				-2.0			

【0235】同表において、所要時間、入力枚数、出力枚数、および処理能力判断手段411の出力は、(表12)と同様である。この処理能力判断手段411の出力に対応して、符号化パターン決定手段410は、上記のように状態遷移をし、符号化パターンを出力する。そして、符号化パラメータ決定手段302は、出力された符号化パターンに対応して、表に示すように1枚のフレーム画像ごとに符号化タイプを出力する。

【0236】同表に示すように、符号化パターン決定手段410の状態は、上記のようにまずS1をとるので、このため表11に示すように符号化パターン「I I I P」が出力される。符号化パラメータ決定手段402はこれに対応して、最初の4つのフレーム画像について「I」「I」「I」「P」を出力する。

【0237】8番目から11番目までの4つのフレーム画像に対しては、符号化パターン決定手段410の状態がS1→S2に遷移したことから、符号化パターン「I P I P」が出力されるので、符号化パラメータ決定手段402の出力は、「I」「P」「I」「P」となる。

【0238】以下、表12に示すように、4つのフレーム画像を処理するごとに、処理能力の判断と、それに対応した符号化パターンの選択がなされ、該パターンに対応した符号化タイプにおいて符号化処理が行われる。

【0239】従来の技術による映像符号化装置では、実施の形態3において(表10)を用いて示したように、符号化結果として得られるフレームレートや、当該符号化装置を構成するハードウェア能力の変動を考慮せずに、符号化タイプ(パターン)、あるいは解像度を決定していたものである。従って、符号化処理の結果として得られるフレームレートが要望される値に近くなるよう設定することが困難であり、ときには不必要な数値となってしまう設定を選定せざるを得ない場合などがあった。これに比べ、本実施の形態4の映像符号化装置においては、実施の形態3と同様に、当該符号化装置の処理能力や符号化結果であるフレームレートの変動を考慮して、指定された解像度に応じて符号化タイプ(パターン)を決定することで、表13と表10との対比において示されるように、指定されたフレームレートに近いフレームレートを実現でき、かつ、より高圧縮率での符号化が実行されていることがわかる。また、実施の形態3においては、処理に要した時間の測定を要するものであったが、本実施の形態4では、実施の形態3のように処理時間を測定が不可能な、または困難な場合であって

も、一時蓄積するデータ量を指標として装置の処理能力を判定することが可能である。

【0240】このように、本実施の形態4による映像符号化装置によれば、符号化手段401、符号化パターン決定手段410を内包した符号化パラメータ決定手段402、処理能力判断手段411、バッファ手段412、および入力フレームレート制御手段413を備えたことで、入力フレームレート制御手段413が入力画像データを定められたレートでバッファ手段412に入力して、入力バッファ手段412は、この入力画像データを一時蓄積し、処理能力判断手段411は、バッファ手段412に蓄積されたデータ量に基づいて、当該符号化装置の処理能力を示す判断結果を出力し、符号化パラメータ決定手段402は、指定されたフレームレートと解像度、そして処理能力判断手段411の出力する判断結果とに対応して符号化パターンを決定して、符号化パラメータを符号化手段401に出力し、符号化手段401はこの符号化パラメータに応じて符号化の処理を行うので、要求される条件を実現しつつ、より高圧縮率の得られる符号化を行うことが可能となる。

【0241】なお、本実施の形態4による映像符号化装置では、指定された解像度に対応して符号化パターンを決定するものとしたが、同様の処理をすることによって、指定された符号化パターン(タイプ)に対応して解像度を決定することも可能であり、要求されるフレームレートと符号化パターンとの下で、より高解像度での符号化処理を行うことが可能となる。

【0242】また、本実施の形態4による装置では、実施の形態3と同様に、符号化の処理を行いながら一時蓄積されるデータ量を検出し、装置の処理能力として取得される符号化の状況に対応して動的に符号化パラメータを変更し得るものである。従って、本実施の形態4による映像符号化装置では、実施の形態1、および2に比較して若干の処理負担は伴うものの、複数の演算処理等を並行して実行する汎用計算機において、映像取り込みに伴った符号化を実行する際などであって、当該計算機装置の状況が変化するような場合でも、その状況の変化に対応して、適切な符号化条件を設定することが可能となるものである。

【0243】もっとも、実施の形態3と同様に、あまり状況の変化がないなど当該符号化装置の処理能力が符号化処理の過程において大きく変動しないと見込まれる場合等には、本実施の形態4による映像符号化装置におい

ても、実施の形態1、および2と同様に、符号化開始に際してパラメータを設定し、以後はその条件で符号化を実施するものとして、制御にかかる処理負担を軽減することも可能である。

【0244】なお、実施の形態1～4では、符号化タイプとして、フレーム内符号化と、順方向予測符号化とを行うものとしたが、符号化タイプはこれに限られるものではない。たとえば、逆方向予測符号化、双方向予測符号化等をもそれぞれ異なる符号化タイプとして採用することができ、さらには、フレーム間予測符号化において動きベクトルの探索範囲を変える場合などについても、異なる符号化タイプとして用いることが可能である。

【0245】また、実施の形態1～4に示した映像符号化方法については、該方法を実行し得る映像符号化プログラムを記録した記録媒体を用いて、パーソナルコンピュータやワークステーション等において、当該プログラムを実行することによって実現できるものである。

【0246】実施の形態5. 本発明の実施の形態5による音声符号化装置は、サンプリングした音声データに対して、変換処理を行うことにより、当該音声符号化装置における処理負担の軽減を図り得るものである。図10は、本発明の実施の形態5による音声符号化装置の構成を示すブロック図、図11は本実施の形態5の符号化装置のハードウェア構成を示す図である。図10に示すように、当該音声符号化装置は、音声入力部501、レジスタ502、入力音声サンプリング部503、音声データ変換部504、帯域分割部505、符号化ビット割り当て部506、量子化部507、符号化部508、および符号化データ記録部509から構成されている。

【0247】音声入力部501は、符号化を行う対象である音声を入力するものである。音声は図11に示すようにマイクロホンから入力されても、あるいはライン入力であっても良い。レジスタ502は、図11のメインメモリまたは外部記憶装置で実現され、符号化処理に用いられる定数を記憶する。入力音声サンプリング部503は、図11のサウンドボード(入力)および制御プログラムによって実現され、音声入力部501が入力した音声に対してサンプリング処理を行う。

【0248】音声データ変換部504は、入力音声サンプリング部503がサンプリング処理したデータに対して、レジスタ502に記憶された定数の値を用いた変換処理をする。帯域分割部505は、音声データ変換部504が変換したデータを帯域分割する。符号化ビット割り当て部506は、帯域分割部505が分割した帯域に対して、符号化ビットを割り当てる。量子化部507は、符号化ビット割り当て部506の割り当てた符号化ビット数に従って、量子化処理を行う。符号化部508は、量子化部507の出力する量子化値を符号化音声データとして出力する。504～508はいずれも、図11のCPU、メインメモリ、およびプログラムで実現さ

れる。符号化データ記録部509は、図11の外部記憶装置、および制御プログラムで実現され、符号化部508から出力された符号化データを当該音声符号化装置の音声符号化処理結果として記録する。

【0249】本実施の形態5において、設定周波数 f_s としては、MPEG Audioで規定される、32kHz、44.1kHz、48kHzの3つのサンプリング周波数のうち、48kHzを採用したものとす。又、変換定数 n は、CPU性能に従って予め定められた値「2」としてレジスタ502に格納されたものであるとする。変換定数 n の値の決定については、装置に用いられるCPUを固定的に想定して、その符号化処理性能に基づいて設定する方法、予めシミュレーションなどによりCPUごとに求めた値の中から、ユーザーによるCPU選択によって選定する方法、符号化処理に先立ちCPUの符号化処理性能を計る演算を行わせ、その結果に基づいて設定する方法などを用いることができる。

【0250】図12は本実施の形態5の符号化装置による音声符号化の動作を示すフローチャート図、図13は本実施の形態5の符号化装置によるサンプリングおよびそれに続く音声データの変換を説明するための図である。以下に本実施の形態5による音声符号化装置による符号化の際の動作を、図12のフローチャートに従って、図10、および13を参照しながら説明する。

【0251】図12のフローのステップ1で、音声入力部501より入力された音声信号は、入力音声サンプリング部503において、設定周波数 f_s をサンプリング周波数としてサンプリングされる。従来例の場合と同様に、このサンプリングはサンプリング周波数 f_s と逆数関係にある時間を t_s として、図13(a)のように行われ、 m 個のサンプリング音声データが出力される。

【0252】ステップ2において、音声データ変換部504は、レジスタ502に記憶された変換定数 n を得て、入力音声サンプリング部503の出力する m 個のサンプリング音声データから $(n-1)$ 個とばしで、計 m/n 個のサンプリング音声データを抽出する。この場合 $(n-1)$ は1であるので、図13(b)に示すように、1個おきに○印をつけたサンプリング音声データが抽出される。そして、音声データ変換部504は抽出したデータが、それぞれ n 個ずつ連続する合計 m 個の変換音声データを作成する。図13(c)のような変換音声データが出力され、この変換音声データは、周波数 f_s の m 個の音声データとなる。

【0253】ここで、ステップ2における音声データ変換を、単に $(n-1)$ 個とばしに抽出するだけで、後の帯域分割符号化に供する場合と比較して考察する。この場合、図13(d)のデータCに示すように、サンプリング周波数 f_s/n の m/n 個の音声データが作成されることとなり、出力される符号化音声データはサンプリング周波数 f_s/n に相当する符号化データにしかなら

ず、サンプリング周波数 f_s としては再生することは出来ない。ここでの想定により、MPEG Audio の規定によって 48 kHz でサンプリングした後、 $n=2$ により、1 個飛ばしに抽出しただけの、 $m/2$ 個の音声データに対して帯域分割符号化を行ったものとする、サンプリング周波数 $f_s/2 = 24 \text{ kHz}$ に相当する符号化音声データしか出力されず、上記 MPEG Audio 規定の 3 つのサンプリング周波数、32 kHz、44.1 kHz、および 48 kHz のいずれによっても再生出来ないことが分かる。

【0254】このため、本実施の形態 5 による符号化処理においては、ステップ 2 の変換処理で、単純に $(n-1)$ 個飛ばしの m/n 個の音声データ (図 13(d) のデータ C) とするだけではなく、同じデータが n 個ずつ連続する変換音声データを作成するものである。図 13(c) のデータ B に示すこの変換音声データは、実質のサンプリング周波数は f_s/n 相当であるが、図 13(b) のデータ A と同様にサンプリング周波数 f_s として扱える m 個の音声データとなる。

【0255】以上のような変換工程であるステップ 2 に続いて、ステップ 3 では、この変換音声データに対して、帯域分割部 505 が M 個の周波数帯域への分割を行う。MPEG Audio の帯域分割では、32 帯域への分割を行う。このステップは第 1 の従来例の場合と同様に行われる。

【0256】ステップ 4 で、符号化ビット割り当て部 506 は、設定周波数 f_s と、変換定数 n とをレジスタ 502 より取得し、これらに基づいて、一般に知られてい*

$$Z_i = C_i \times X_i \quad (i = 0 \sim 511)$$

$$Y_i = \sum_{j=0}^7 Z_{64+j} \quad (i = 0 \sim 63)$$

$$S_i = \sum_{k=0}^{63} Y_k \cos((2i+1)(k-6) \pi / 64) \quad (i = 0 \sim 31) \quad (3)$$

【0261】ただし、

X_i : 入力音声データ

S_i : 帯域分割後の音声データ

また、係数 C_i は、MPEG オーディオの規格による、サンプル番号と係数とを対比させた係数表より得られる。

【0262】ここで、式 (1) および (2) は、通常 m 個の音声データに対して演算されるが、ステップ 2 において変換した、抽出したサンプリング音声データが n 個ずつ連続する変換音声データに対して演算を行う場合、この n 個ずつのサンプリング音声データは同じものである、これらを連続的に扱う部分については、 m/n 個の音声データに対して演算すればよく、式 (1) に

* サンプリング定理より再生可能な限界となる制限周波数 $f_s/2n$ を算出する。そして、 M 個に分割された帯域の中で、再生可能である制限周波数より小さい周波数の帯域には符号化ビットを割り当て、再生可能でない制限周波数より大きい周波数の帯域には符号化ビットを割り当てないものとして、符号化ビット割り当て数を決定する。符号化ビット割り当て数は符号化ビット割り当て部 506 から量子化部 507 に伝えられる。

【0257】ステップ 5 において、量子化部 507 は、符号化ビット割り当て数に従って、それぞれの帯域ごとに音声データを量子化して量子化値を出力し、ステップ 6 で符号化部 508 は、この量子化値により、符号化音声データを出力する。出力された符号化音声データは、符号化データ記録部 509 に記録される。図 12 のフローに示すように、以上の過程は、符号化を行う対象の音声の入力が続く間繰り返され、音声の入力終了後、速やかに符号化は完了する。

【0258】本実施の形態 5 による音声符号化装置の効果について、ステップ 3 の帯域分割において、ステップ 2 で得られた変換音声データを用いることによる処理量の軽減を、MPEG Audio での帯域分割方法において考察する。

【0259】ここで想定している MPEG Audio の帯域分割では、32 帯域への分割のため以下の演算を行う。

【0260】

【数 1】

(1)

(2)

(3)

については、演算量は $1/n$ にまで軽減され、式 (2) についても $1/n$ に軽減することもできる。

【0263】ここでは、 $n=4$ の場合について説明を行う。この場合、式 (1) では、 $X_0 \sim X_3$ の音声データは、4 個の X_0 が連続するものである、 $X_0 = X_1 = X_2 = X_3$ となり、また、 C_i についても $C_0 = C_1 = C_2 = C_3$ となる一つの値を用いて演算を行うことができる。結局 Z_0, Z_1, Z_2, Z_3 の 4 つの値は 1 回の演算によって求められることとなり、式 (1) では $1/4$ の演算量において全ての Z_i が求められる。なお、式

(1) の演算において、 $C_0 = C_1 = C_2 = \dots = C_n$ となる一つの値で代表させるには、 $C_0 \sim C_n$ のいずれかの値を用いる方法や $C_0 \sim C_n$ の平均値を用いる方法な

どが用いられる。

【0264】次にこの Z_i を用いた式(2)でも、0から64個飛びの値を8回加算しているだけとなり、 n が2のべき乗の場合は、 Y_i は n 個ずつ同じ値となり、演算は $1/n$ に削減される。ただし、 n が2のべき乗でない場合、例えば $n=3$ の場合では、 Y_0 を求めるために加算する Z_{64} と、 Y_1 を求めるために加算する Z_{65} とは等しくないため、 $Y_0=Y_1$ とはならず、結局式(2)の演算量は削減されないこととなる。

【0265】ただし、この n が2のべき乗でない場合について、設定により演算量を削減することは可能である。例えば、 $n=3$ の場合、 $X_0=X_1=X_2$ 、 $X_3=X_4=X_5$ 、…、 $X_{27}=X_{28}=X_{29}$ 、 $X_{30}=X_{31}$ 、となるように、32帯域への分割において、 $i=0\sim31$ の間で同じ値がなるべく多く連続するように、ステップ2における音声データ変換を行う。つまり、最後の2個を同じとして、それ以外は同じ値の連続する3つ組が並ぶ形式の音声データ列に変換しておき、この変換音声データを用いて、ステップ3の帯域分割の際の式(1)および式(2)の演算を行えば、 $1/3$ に近い削減ができる。式(3)については、音声データ変換後も演算量は変わらない。

【0266】また、ステップ4における符号化ビット割り当ての際に、制限周波数 $f_s/2n$ 以下の帯域のみ符号化ビットを割り当てるのは、元々サンプリング周波数 f_s でサンプリングした音声データを、ステップ2の音声データ変換に際して $(n-1)$ 個飛ばしに抽出することが、サンプリング周波数 f_s/n でサンプリングしたことと同等になり、公知のサンプリング定理より、 $f_s/2n$ 以上の周波数帯域は再生出来ないことが分かるので、再生可能な $f_s/2n$ 以下の帯域のみを符号化の対象とするためである。制限周波数以上の帯域には符号化ビットを割り当てず、その帯域に対しての量子化が不要なことから、ステップ5における量子化処理は、 $1/2n$ に負担が軽減される。

【0267】このように、本実施の形態5の音声符号化装置によれば、レジスタ502と、音声データ変換部504とを備えたことで、入力音声サンプリング部503が設定周波数によりサンプリング処理した m 個のサンプリング音声データに対して、音声データ変換部504が、レジスタ502に記憶した変換定数 n に基づいて、 $n-1$ 個とばしにサンプリング音声データを合計 m/n 個抽出し、この抽出したサンプリング音声データがそれぞれ n 個ずつ連続する m 個の変換音声データとすることで、これに続く帯域分割部505による処理に際し、演算量を大きく削減することができ、しかも、単に抽出によって音声データを減らした場合とは異なり、元の設定周波数において再生可能な符号化データを得ることができる。又、帯域分割した帯域に対し、符号化ビット割り当て部506が、サンプリング定理によって再生可能で

ない帯域には符号化ビットを割り当てず、その帯域に対しては量子化部507による量子化が不要なことから、量子化処理の負担が $1/2n$ に軽減される。従って、CPU性能不足などにより、従来の方法では、音声入力に伴った実時間符号化処理が困難または不可能な場合にも、定数の設定により負担を軽減することで音声符号化処理を実時間で行うことが可能となる。

【0268】なお、本実施の形態5では、変換音声データの作成においてサンプリング音声データが並ぶ形式のものを作成したが、サンプリング音声データの間に両側の音声データを平均した音声データなど、適当な音声データを $n-1$ 個挿入して、同様の効果を得ることも可能である。

【0269】実施の形態6. 本発明の実施の形態6による音声符号化装置は、実施の形態5と同様にサンプリングした音声データに対して、変換処理を行うことにより、当該音声符号化装置における処理負担の軽減を図り得るものであるが、データ量削減を音声データ変換処理でなくサンプリング処理において行うものであることが、実施の形態5とは異なる。

【0270】図14は、本発明の実施の形態6による音声符号化装置の構成を示すブロック図である。同図に示すように、当該音声符号化装置は、音声入力部601、レジスタ602、入力音声サンプリング部603、音声データ変換部604、帯域分割部605、符号化ビット割り当て部606、量子化部607、符号化部608、および符号化データ記録部609から構成されている。また、本実施の形態6の符号化装置のハードウェア構成も図11に示される実施の形態5のものと同様である。

【0271】入力音声サンプリング部603は、実施の形態1の入力音声サンプリング部503とは異なり、設定周波数をそのままサンプリング周波数とするのではなく、レジスタ602から変換定数を得て、設定周波数とこの変換定数とを用いて定められるサンプリング周波数を用いてサンプリング処理を行う。また、音声データ変換部604は、実施の形態1の音声データ変換部104とは異なり、サンプリング音声データの抽出を行わず、音声データ挿入のみを行って、変換音声データを作成する。音声入力部601、帯域分割部605、符号化ビット割り当て部606、量子化部607、符号化部608、および符号化データ記録部609は実施の形態5における501、および505～509と同様である。図15は本実施の形態6の符号化装置による音声符号化の動作を示すフローチャート図である。また、本実施の形態6においても、サンプリングと音声データ変換処理の説明には図13を用いる。以下に本実施の形態6による音声符号化装置による符号化の際の動作を、図15のフローチャートに従って、図14を参照しながら説明する。実施の形態5の場合と同様に、設定周波数 f_s は、MPEG Audioで規定される48kHz、変換定数

n は、「2」とする。

【0272】図15のフローのステップ1において、入力音声サンプリング部603は設定周波数 f_s と、変換定数 n とをレジスタ602より得て、これらから、実行サンプリング周波数 f_s/n を決定し、音声入力部601より入力された音声信号は、入力音声サンプリング部603において実行サンプリング周波数 f_s/n によりサンプリングされる。このサンプリングの結果、図13(d)のデータCのような、 m/n 個のサンプリング音声データが出力される。

【0273】ステップ2において、音声データ変換部604は、レジスタ602より変換定数 n を得て、入力音声サンプリング部603の出力する m/n 個のサンプリング音声データに基づいて、各サンプリング音声データが n 個ずつ連続する m 個の変換音声データを作成する。図13(c)のデータBのような変換音声データが出力され、この変換音声データは、周波数 f_s の m 個の音声データとなる。

【0274】実施の形態5において説明したように、図13(d)のデータCの音声データはサンプリング周期 f_s で再生できないが、変換を行って図13(c)のデータBのような音声データとすることで、サンプリング周期 f_s で再生可能な符号化データを得ることができる。

【0275】ステップ2において得られた変換音声データは、実施の形態5の場合のステップ2で得られた変換音声データと同等のものとなるので、これ以降のステップ3～6は実施の形態5におけるステップ3～6と同様に実行される。そして、ステップ1～6は、音声の入力が続く間繰り返され、音声の入力終了後、速やかに符号化は完了する。

【0276】本実施の形態6の音声符号化装置においても、帯域分割の段階と、量子化の段階とにおいて、実施の形態5の装置で説明したのと同様の演算作業の削減が可能となり、CPU性能等に応じたレベルで、音声入力に伴っての実時間符号化処理が可能となる。

【0277】このように、本実施の形態6の音声符号化装置によれば、レジスタ602、入力音声サンプリング部603、および音声データ変換部604を備えたことで、入力音声サンプリング部603が設定周波数 f_s とレジスタ602に記憶した変換定数 n とを用いて実行サンプリング周波数 f_s/n を定めてサンプリング処理を行い、得られた m/n 個のサンプリング音声データに対して、音声データ変換部604が音声データの挿入を行うことによって、 m 個の音声データからなる変換音声データを得ることで、実施の形態5と同様に、装置の処理負担の軽減を図ることが可能となる。加えて本実施の形態6の符号化装置では、サンプリング周波数 f_s/n でサンプリングを行うことにより、サンプリング入力において音声データを一時的に保存するバッファメモリなどは実施の形態5による装置の場合の $1/n$ の容量でよ

く、また、サンプリング周波数の上限が f_s まで無いようなサウンドボードを用いる場合でも、動作可能であるという利点を持ち、より少ないハードウェア資源においても、装置資源を活用して、音声入力に伴っての実時間符号化処理を行うことが可能となる。

【0278】なお、本実施の形態6においても実施の形態5と同様、変換音声データの作成においてはサンプリング音声データが並ぶ形式のものを作成したが、適当な音声データを $n-1$ 個挿入して、同様の効果を得ることも可能である。

【0279】実施の形態7. 本発明の実施の形態7による音声符号化装置は、入力されるデータの量に対応して変換定数を変更することにより、状況に応じた符号化を実行できるように図るものである。

【0280】図16は、本発明の実施の形態7による音声符号化装置の構成を示すブロック図である。同図に示すように、当該音声符号化装置は、音声入力部701、レジスタ702、入力音声サンプリング部703、音声データ変換部704、帯域分割部705、符号化ビット割り当て部706、量子化部707、符号化部708、符号化データ記録部709、入力バッファ7010、および入力バッファ監視部7011から構成されている。この構成は、実施の形態5による音声符号化装置に入力バッファ7010と、入力バッファ監視部7011とを追加した構成である。また、本実施の形態7の符号化装置のハードウェア構成も図11に示される実施の形態5のものと同様である。

【0281】入力バッファ7010は、主としてメインメモリ等のメモリで実現され、データを一時記憶する。入力バッファ監視部7011は、CPU、メインメモリ、およびプログラムで実現され、入力バッファ7010に一時記憶のため保持されるデータ量を調べて、このデータ量を予め設定された値と比較し、その結果によって、レジスタ702の変換定数 n の値を変更する。レジスタ702は、記憶する変換定数の値が入力バッファ監視部7011によって変更されることを除いて実施の形態5のレジスタ502と同様である。入力音声サンプリング部703は、サンプリング音声データを入力バッファ7010に出力する点を除いて、実施の形態5の入力音声サンプリング部503と同様である。音声データ変換部704は、入力バッファ7010からサンプリング音声データを取りだして処理対象とする点を除いて、実施の形態5の音声データ変換部504と同様である。また、音声入力部701、帯域分割部705、符号化ビット割り当て部706、量子化部707、符号化部708、および符号化データ記録部709は実施の形態5における501、および505～509と同様である。

【0282】図17は本実施の形態7の符号化装置による音声符号化の動作を示すフローチャート図である。以下に本実施の形態7による音声符号化装置による符号化

の際の動作を、図17に従って、図16を参照しながら説明する。実施の形態5の場合と同様に、設定周波数 f_s は、MPEG Audioで規定される48kHzとする。又、変換定数 n は、CPU性能に従って予め定められた値「1」が初期値としてレジスタ702に格納されているものとする。

【0283】図17のフローのステップ1で、音声入力部701より入力された音声信号は、入力音声サンプリング部703において、実施の形態5と同様にサンプリングされ、ステップ2において、サンプリング音声データは入力バッファ7010に書き込まれて一時記憶される。ステップ3において、音声データ変換部704は入力バッファ7010より、一時記憶されたサンプリング音声データを読み出す。そして、後述するステップ4の後に実行されるステップ5における音声データの変換以降、ステップ9の符号化データ出力までは、実施の形態5における図12のフローのステップ2～6と同様に実行されるので、ステップ5～9における動作については説明を省略する。

【0284】ステップ3が実行された後、ステップ4では、入力バッファ監視部7011が、入力バッファ7010に保持されたデータ量を調べて、このデータ量を予め設定した値と比較し、比較結果に基づいて、レジスタ702に記憶される変換定数 n の値を変更する。入力バッファ7010を監視して、変換定数 n の値を制御するには様々な方法が採用できるが、ここでは以下の様に行われるものとする。

【0285】CPUの負担増大などにより、当初の設定では、音声入力にともなう符号化処理ができなくなった場合、入力バッファ7010については、書き込みは同じペースで行われるのに対して、符号化処理のための読み出しのペースが落ちるため、データ量は増大する。

【0286】入力バッファ監視部7011は、入力バッファ7010のデータ量が、予め設定したバッファフルレベルBFを越えた場合は、現状の設定での実時間符号化処理が不可能であると判断し、レジスタ702に記憶される変換定数 n の値を、1だけ増加させて $n=2$ に変更する。それ以後のフロー図のステップ5～9においては、ステップ5ではデータを1個飛ばしに間引いて、2個同じデータが続く形式に変換し、これをステップ6で帯域分割することにより、ステップ6の帯域分割での一部の処理を1/2に軽減させる。又、ステップ7では、各帯域に符号化ビットの割り当てを行う際、周波数 $f_s/4$ 以下の帯域に対してのみ符号化ビット割り当てることにより、ステップ8での量子化処理を1/4に軽減させる。このようにして、入力バッファ監視部7011は変換定数 n の値を変更することにより、CPUに対する負担の軽減を図る。

【0287】図17のフローにおける繰り返しで、ステ

ップ4において、なおも入力バッファ7010のデータ量がバッファフルレベルBFを越える場合、入力バッファ監視部7011は、レジスタ2の変換定数 n を変更し、さらに1増加して $n=3$ とする。これにより、ステップ5では、データを2個飛ばしに間引いて、3個同じデータが続く形式に変換することにより、ステップ6の帯域分割での一部の処理を1/3に軽減させ、ステップ7で各帯域に符号化ビットの割り当てを行う際、周波数 $f_s/6$ 以下の帯域に対してのみ符号化ビットを割り当てることにより、ステップ8での量子化処理を1/6に軽減させる。以後、ステップ4で入力バッファ7010のデータ量がバッファフルレベルBF以下になるまで、入力バッファ監視部7011は、レジスタ2の n の値を増加させる。

【0288】逆に、ステップ4において、入力バッファ7010の保持するデータ量が予め設定したバッファエンブティレベルBEを下回る場合は、入力バッファ監視部7011は、符号化処理能力に余力があると判断する。なるべく変換定数 n の値が少ない方が、音声データの間引きと、高周波成分のカットがなく、高品質の符号化データが得られるので、入力バッファ監視部7011は、変換定数 n の値を1だけ減少させ、以後は上記と同様に入力バッファ7010のデータ量がバッファエンブティレベルBE以上になるまで、図17のフローの繰り返しにおいて、ステップ4でレジスタ702の記憶する変換定数 n の値を1ずつ減少させる。

【0289】なお、上記の方法では、変換定数 n の値を制御するために、バッファフルレベルBFとバッファエンブティレベルBEの2つの値を用いたが、バッファフルレベルBFのみを用いてもよく、この場合、入力バッファのデータ量が予め設定したバッファフルレベルBFに達するまで変換定数 n の値を増加し、音声入力と符号化処理とが釣りあうとき、すなわちデータ量がBFに達したときに、変換定数 n を増加するのを止める様に制御を行う。

【0290】このように、本実施の形態7の音声符号化装置によれば、実施の形態5による音声符号化装置に、入力バッファ7010と、入力バッファ監視部7011とを追加する構成としたことで、サンプリング音声データをこの入力バッファ7010に一時記憶した後に読み出して、それ以後の処理を行うものとし、また、入力バッファ監視部7011が、入力バッファ7010の保持するデータ量を調べることににより、これをその時点におけるCPUの符号化処理能力の指標として、レジスタ702に記憶する変換定数 n の値を状況に応じて動的に制御することによって、CPUがその時点で符号化処理可能な、最も高品質な音声符号化を行うように図ることが可能となる。

【0291】実施の形態8. 本発明の実施の形態8による音声符号化装置は、出力されるデータの量に対応して

変換定数を変更することにより、状況に応じた符号化を実行できるように図るものである。図18は、本発明の実施の形態8による音声符号化装置の構成を示すブロック図である。同図に示すように、当該音声符号化装置は、音声入力部801、レジスタ802、入力音声サンプリング部803、音声データ変換部804、帯域分割部805、符号化ビット割り当て部806、量子化部807、符号化部808、符号化データ記録部809、および符号化データ監視部8012から構成されている。この構成は、実施の形態5による音声符号化装置に符号化データ監視部8012を追加した構成である。また、本実施の形態8の符号化装置のハードウェア構成も図11に示される実施の形態5のものと同様である。

【0292】符号化データ監視部8012は、CPU、メインメモリ、およびプログラムで実現され、符号化部808より出力される単位時間当たりの符号化データ量を調べて、このデータ量を予め設定された値と比較し、その結果によって、レジスタ802の変換定数 n の値を変更する。レジスタ802は、記憶する変換定数の値が符号化データ監視部8012によって変更されることを除いて実施の形態5のレジスタ502と同様である。音声入力部801、入力音声サンプリング部803、音声データ変換部804、帯域分割部805、符号化ビット割り当て部806、量子化部807、符号化部808、および符号化データ記録部809は実施の形態5における501、および503～509と同様である。

【0293】図19は本実施の形態8の符号化装置による音声符号化の動作を示すフローチャート図である。以下に本実施の形態8による音声符号化装置による符号化の際の動作を、図19に従って、図18を参照しながら説明する。実施の形態5の場合と同様に、サンプリング周波数 f_s は、MPEG Audioで規定される48kHzとする。又、変換定数 n は、CPU性能に従って予め定められた値「1」が初期値としてレジスタ802に格納されているものとする。図19のフローのステップ1からステップ6までは、実施の形態5におけるステップ1～6と同様に実行される。そして、ステップ7では、符号化データ監視部8012が、符号化部808より出力される単位時間当たりの符号化データ量を調べて、このデータ量を予め設定された値と比較し、その結果によって、レジスタ802の変換定数 n の値を変更する。符号化データの量を監視して変換定数 n の値を制御するには様々な方法が採用できるが、ここでは以下の方法に従って行われるものとする。

【0294】CPUの負担増大などにより、当初の設定では、符号化処理が間に合わなくなった場合、符号化処理のペースが落ちるため、出力される符号化データ量は減少する。ステップ7において、符号化データ量が、予め設定された符号化最低レベル CL に達しない場合は、符号化データ監視部8012は、実施の形態7に示した

入力バッファ監視部7011と同様、レジスタ802の変換定数 n の値を増加させることにより、CPUの負担を軽減させるよう図る。図19の処理を繰り返し、ステップ7において、単位時間当たりの符号化処理量が符号化最高レベル CH を下回らない場合は、高品質の符号化が行えるよう、レジスタ802の変換定数 n の値を減少させることも実施の形態7と同様である。実施の形態7における入力バッファ監視部7011による制御と同様に、本実施の形態8における符号化データ監視部8012も、符号化データ量が適切と判定されるまでは、変換定数 n の値を変更し続ける。又、このように符号化最低レベル CL と符号化最高レベル CH の2つの値を用いず、符号化最低レベル CL のみを用いても制御可能である点についても、実施の形態7と同様である。

【0295】このように、本実施の形態8の音声符号化装置によれば、実施の形態5による音声符号化装置に、符号化データ監視部8012を追加する構成としたことで、符号化データ監視部8012が、単位時間当たり出力される符号化データ量を調べることにより、これをその時点におけるCPUの符号化処理能力の指標として、レジスタに802に記憶する変換定数 n の値を状況に応じて動的に制御することによって、CPUがその時点で符号化処理可能な、最も高品質な音声符号化を行うように図ることが可能となる。

【0296】なお、実施の形態7および8については、実施の形態5に準じたものとして、入力音声サンプリング部がサンプリング周波数 f_s によるサンプリングで m 個のサンプリング音声データとし、次に音声データ変換部が、 $(n-1)$ 個とばしの間引きを行うものとしたが、実施の形態6の装置の場合のように、入力音声サンプリング部がサンプリング周波数を f_s/n としてサンプリングし、 m/n 個のサンプリング音声データを得て、これを音声データ変換部が変換して m 個の変換音声データを得る方式としてもさしつかえなく、ソフトウェア上の設定変更で容易に行える。またその場合、実施の形態6で説明したように、バッファメモリの容量低減や、サンプリング周波数の制限の厳しいサウンドボードの使用も可能、といった効果は、同様に得られる。

【0297】なお、実施の形態5～8による符号化においては、実質的にオーディオデータの間引きや高周波成分の除去を行うため、それに伴い音質が劣化することにはなる。しかしその場合でも、性能の低いCPUによっても、ハードウェア的追加等を要せずソフトウェア的に、MPEG Audioなどの帯域分割符号化データを実時間で作成でき、これを、動画符号化の国際標準として広く用いられるMPEGデータとして利用することが可能となる。また、変換定数の値を調整することで、CPUの符号化処理性能にあわせて、間引き具合や除去する高周波成分の割合を制御できるため、高性能なCPUのみならず性能が不十分なCPUでもその符号化処理能

力なりの音質で符号化することが出来、幅広い性能レベルのCPUで符号化処理が実現できる。但し、ハードウェア面に関しては、CPUが高性能であるほど、またサウンドボードの機能や装置内でのデータ伝送速度が高いほど、高品質な符号化が可能である。

【0298】また、実施の形態5～8の音声符号化は、音声符号化制御プログラムとして記録媒体に記録し、パーソナルコンピュータ、ワークステーションその他の装置において実行することが可能である。また、実施の形態5～8では、符号化データを記憶装置に保存することとしたが、ネットワーク等を介して他の機器に伝達し、他の機器において記録または利用することも可能である。また、実施の形態5～8では、CPU処理によるものとして説明したが、CPUの代わりにDSPを用いたソフトウェア処理によっても、同様である。

【0299】実施の形態9. 本発明の実施の形態9による音声符号化装置は、単位期間に区切られたサンプリングデータに対して、設定された定数に応じて、単位期間ごとにデータ処理を行うか否かを制御することで、処理負担の軽減を図り得るものである。図20は、本発明の実施の形態9による音声符号化装置の構成を示すブロック図である。図21は本実施の形態9による音声符号化のフローチャート図、図22は本実施の形態9による音声符号化を説明するための概念図である。また、本実施の形態6の符号化装置のハードウェア構成も実施の形態5のものと同様であり、説明には図11を用いる。

【0300】図20に示すように、本実施の形態9による音声符号化装置は、音声入力部901、レジスタ902、入力音声サンプリング部903、判定制御部(単位期間判定)904、帯域分割部905、符号化ビット割り当て部906、量子化部907、符号化部908、符号化データ記録部909、および固定的符号レジスタ910から構成されている。

【0301】音声入力部901は、符号化を行う音声を入力するものである。音声は図11に示すようにマイクロホンから入力されても、あるいはライン入力であっても良い。単位期間判定定数レジスタ902は、図11のメインメモリまたは外部記憶装置で実現され、単位期間判定定数を記憶する。入力音声サンプリング部903は、図11のサウンドボード(入力)および制御プログラムによって実現され、音声入力部901が入力した音声に対してサンプリング処理を行う。判定制御部904は、入力音声サンプリング部903がサンプリング処理したデータに対して、レジスタ902に記憶された定数の値を用いて符号化対象期間であるか否かを判断する。帯域分割部905は、判定制御部904により符号化対象期間であるとされた場合のみ、サンプリングデータを帯域分割する。符号化ビット割り当て部906は、帯域分割部905が分割した帯域に対して、符号化ビットを割り当てる。量子化部907は、符号化ビット割り当て

部906の割り当てた符号化ビット数に従って、量子化処理を行う。符号化部908は、量子化部907の出力する量子化値を符号化音声データとして出力する。本実施の形態9では符号化部908は、判定制御部904において、符号化対象期間でないと判断された場合には、後述する固定的符号レジスタ910に記憶された、帯域出力ゼロに相当する符号化データdNを符号化音声データとして出力する。904～908はいずれも、図11のCPU、メインメモリ、およびプログラムで実現される。符号化データ記録部909は、図11の外部記憶装置および制御プログラムで実現され、出力された符号化データを記録する。固定的符号レジスタ910は、図11のメインメモリまたは外部記憶装置で実現され、帯域出力ゼロに相当する符号化データdNを記憶する。このように構成される本実施の形態9による音声符号化装置による符号化の際の動作を、以下に図21のフローチャートに従って、図20、および22を参照しながら説明する。

【0302】図21のフローのステップ1で、音声入力部901より入力された音声信号は、入力音声サンプリング部903において、設定された周波数 f_s をサンプリング周波数としてサンプリングされる。これにより、周波数 f_s のサンプリングデータが判定制御部904に出力される。

【0303】ステップ2において判定制御部904は、上記サンプリングデータについて、符号化対象期間であるか否かの判定を行う。この判定では、まず、1回の帯域分割で対象とする入力音声サンプル数 p に相当する期間を単位期間 t_i とし、単位期間ごとに符号化対象期間であるか否かの判定を実行する。また、判定に用いられる単位期間判定定数 k は1以上の整数としてあらかじめシステムによって設定され、レジスタ902に記憶される定数である。判定は、単位期間 t_i について、任意の整数 n につき、 $i = n \times k + 1$ が成立するとき符号化対象期間であるとし、成立しないときは符号化対象期間でないとするものとして行われる。

【0304】ステップ2の判定で単位期間 t_i が符号化対象期間の場合は、ステップ3～6が実行されることにより、従来例と同様の処理が行われる。すなわち、まずステップ3において、単位期間 t_i の音声データに対して、帯域分割部905が M 個の周波数帯域に分割を行う。このステップは、図59および図60を用いて説明した第1の従来例の場合と同様に行われる。ステップ4では、符号化ビット割り当て部906で、各帯域に対して符号化ビット数が割り当てられ、その割り当てが量子化部907に伝えられる。ステップ5において、量子化部907は、符号化ビット割り当て数に従って、帯域分割部905が分割したそれぞれの帯域ごとの単位期間 t_i の音声データを量子化して量子化値を出力する。そして、ステップ6では、符号化部908が、量子化部9

07の出力である量子化値により符号化音声データを構成して出力し、符号化音声データは符号化データ記録部909において記録される。

【0305】一方、ステップ2の判定において、単位期間 t_i が符号化対象期間でない場合は、ステップ3～6の帯域分割、符号化ビット割り当て、および量子化は行われず、ステップ2に続いてステップ7が実行される。ステップ7では、符号化部908は、固定的符号レジスタ910から固定的符号化データ d_N を取得して、これを符号化データとして出力を行う。ここで固定的符号化データ d_N は、帯域分割における各帯域の出力をゼロとして、固定的符号レジスタ910に予め設定されたデータである。出力された符号化音声データは、符号化記録部909において記録される。図21のフローに示すように、以上の過程は、符号化を行う対象の音声の入力が続く間繰り返され、音声の入力終了後、速やかに符号化は完了する。

【0306】入力音声サンプル数 $p=32$ で、変数定数 $k=3$ に設定された場合を想定し、図22の概念図を用いてさらに説明する。図に示すように、音声データが入力され、最初の単位期間 t_1 では、 $i=1=0 \times 3 + 1$ が成立する($n=0$)ので符号化対象期間であり、この単位期間分のサンプリングデータは32個の帯域信号に分割され、量子化、符号化されて、符号化データ d_1 が出力される。続く単位期間 t_2 および t_3 では、 $i=n \times k + 1$ を満たす整数 n はなく、符号化対象期間でないと判定されるので、上記一連の処理はされることなく、固定的な符号化データ d_N が出力される。固定的符号化データ d_N については、上記のように32個の帯域信号ゼロとして予め設定されたデータである。この後、単位期間 t_4 では、 $i=4=1 \times 3 + 1$ が成立する($n=1$)ので符号化対象期間であり、この単位期間分のサンプリングデータは t_1 のデータと同様に、帯域分割、量子化、符号化されて、符号化データ d_4 が出力される。以下同様の処理となる。

【0307】本実施の形態9による音声符号化では、上記のように、入力音声に基づく符号化データ d_1 、および d_4 の間に、出力ゼロのデータ d_N が $(k-1)$ 個入った符号化データが得られることとなる。第1の従来例において説明したように、MPEG1 Audioのレイヤ1の音声符号化の場合、入力音声サンプルとしては対象とする32サンプルを中心に前後512サンプルを用いて、32帯域へ分割し帯域ごとの音声データを出力する。従ってこの32サンプル分の期間に相当する帯域出力を出力ゼロとして符号化したのち復号再生したとしても、該部分で音が途切れるわけではなく、その前後の帯域出力の符号化データとともに復号再生される。従って、再生音声のエンベロープ(音声の時間的変化)は連続したものとなるので、人間の聴覚では、あまり大きな音質劣化を感じることはない。

【0308】このように、本実施の形態9による音声符号化装置においては、単位期間判定定数 k を記憶するレジスタ902と、単位期間判定定数 k に基づいて、単位期間分のサンプリングデータにつき、そのデータが符号化対象期間のものであるかどうかを判定する判定制御部904と、符号化対象期間以外のデータを処理して得られる符号化データの代替に用いる固定的符号化データを記憶する固定的符号レジスタ910とを備えたことで、サンプリングデータのうち、符号化対象期間に属する $1/k$ のサンプリングデータのみに対して帯域分割以降の処理を行い、符号化対象期間に属さない残りのサンプリングデータについては該処理を行わず、その分帯域分割出力ゼロとして、固定的な符号化データ d_N を出力するので、 k の値の設定により、帯域分割、符号化ビット割り当て、量子化、および符号化データ生成の各処理について、それぞれの段階での処理量 $1/k$ に負担を軽減することが可能となる。

【0309】従って、CPUの性能不足などにより、従来の方法では、音声入力に伴った実時間符号化処理が困難または不可能な場合にも、変数定数 k の設定により負担を軽減することで音声符号化処理を実時間で行うことが可能となる。変数定数 k の値の決定については、装置に用いられるCPUを固定的に想定して、その符号化処理性能に基づいて設定する方法、予めシミュレーションなどによりCPUごとに求めた値の中から、ユーザーによるCPU選択によって選定する方法、符号化処理に先立ちCPUの符号化処理性能を計る演算を行わせ、その結果に基づいて設定する方法などを用いることができる。

【0310】なお、前述のように帯域出力ゼロの期間があっても再生の際に音がとぎれることはないが、この期間が長いほど音の劣化が大きくなるため、単位期間判定定数の設定については、帯域出力ゼロの期間を32サンプルとするように、つまり $k=2$ に留めるのが望ましい。ただし、その音の劣化を容認しても符号化処理を実行させる必要があれば、 k を大きくすることにより、装置の能力に応じた実時間符号化は可能である。

【0311】実施の形態10. 本発明の実施の形態10による音声符号化装置は、帯域分割に際し一部の演算処理を省略することで、処理負担の軽減を図り得るものである。図23は、本発明の実施の形態10による音声符号化装置の構成を示すブロック図である。図23に示すように、本実施の形態10による音声符号化装置は、音声入力部1001、入力音声サンプリング部1003、帯域分割部1005、符号化ビット割り当て部1006、量子化部1007、符号化部1008、符号化データ記録部1009、およびレジスタ1011から構成されている。また、本実施の形態10の装置も実施の形態9と同様、図11に示されるハードウェア構成である。

【0312】同図において、レジスタ1011は、メイ

ンメモリまたは外部記憶装置で実現され、帯域分割処理における演算実行制御に用いられる演算処理判定定数を記憶する。又、本実施の形態10の帯域分割部1005は、レジスタ1011より演算処理判定定数の値を得て、帯域分割における演算処理を中止する演算処理中止部を内包したものである。音声入力部1001、入力音声サンプリング部1003、符号化ビット割り当て部1006、量子化部1007、符号化部1008、および符号化データ記録部1009は、実施の形態9の901、903、および906~909と同様であり、説明を省略する。

【0313】図24は本実施の形態10の符号化装置による音声符号化のフローチャート図、図25は帯域分割に際して基本低域通過フィルタ処理の演算に用いられる係数を示す図である。以下に本実施の形態10による音声符号化装置の動作を、図24に従って、図23、およ*

$$Z_i = C_i \times X_i \quad (i = 0 \sim 511)$$

$$Y_i = \sum_{j=0}^7 Z_{64+j} \quad (i = 0 \sim 63)$$

$$S_i = \sum_{k=0}^{63} Y_k \cos((2i+1)(k-6)\pi/64) \quad (i = 0 \sim 31) \quad (3)$$

【0317】この中で式(1)に注目してみると、512個の($i = 0 \sim 511$)入力音声データ X_i に対して、表より求められる数 C_i による乗算処理がされるものである。係数 C_i は、MPEGオーディオの規格による、サンプル番号と係数とを対比させた係数表より得られるが、これをグラフ化すると図25のようになり、両端ほど0に近づくことが分かる。そして、式(1)は乗算処理であるため、係数 C_i が0に近づけば、積 Z_i も0に近づくこととなる。さらに式(2)では、式(1)で求めた Z_i を加算するだけなので、係数 C_i が0に近い項については寄与は小さく、かかる項については Z_i を求めて加算する必要性は低い。

【0318】このことから、式(1)については、係数 C_i が0に近い、つまり i が0または511に近い分について Z_i を求める演算を行わず $Z_i = 0$ とし、又、式(2)でも $Z_i = 0$ に相当する項は加算することなく、帯域分割の演算を行えば、帯域分割の精度を幾分損じることとはなるものの、演算量の軽減が図れる。この場合、図25からも分かるように、 C_i は、32個を単位として変化しており、上記演算を打ち切る間隔32個単位で決定することが望ましい。従って、演算を実行する演算対象区間は、 $i = 32q$ から $i = 32(8-q) + 255$ の区間として表すことができる。ここで、演算処理判定定数 q は、 $0 \leq q \leq 7$ を満たす整数として、装置性能等に応じて予め設定してレジスタ1011に記憶さ

*び図25を参照しながら説明する。

【0314】図24のフローのステップ1は実施の形態9と同様に実行され、音声入力部1001より入力された音声信号が、入力音声サンプリング部1003においてサンプリングされ、サンプリングデータが得られる。

【0315】続いてステップ2では帯域分割部1005により、サンプリングデータに対する帯域分割処理が行われる。実施の形態9で説明したように、MPEG1 Audioの帯域分割では、対象とする32サンプルを中心に前後512サンプルを用いて、32帯域へ分割し帯域ごとの音声データを得るものであり、このため基本低域通過フィルタ処理を行う。この基本低域通過フィルタ処理では、以下に示す式(1)(2)(3)の演算を行う。

【0316】

【数2】

(1)

(2)

(3)

せておくものである。上記のように演算対象区間を制限することにより、 $q \times 1/8 = q/8$ だけ演算処理を省くことが可能となるので、演算処理判定定数 q を大きくするほど、演算のための処理負担を軽減することができる。図24のフローのステップ2で上記のように一部の演算を省略して帯域分割が行われた後は、この帯域信号が処理の対象となり、ステップ3以降については、第1の従来例のステップ3以降と同様であるので、説明を省略する。

【0319】このように、実施の形態10による音声符号化装置においては、演算処理判定定数 q を記憶するレジスタ1011を備え、帯域分割部1005は、演算処理判定定数 q に基づき、帯域分割処理における基本低域通過フィルタの演算処理を一部のサンプルについて省略するものとしたことで、演算処理判定定数 q の値を制御することにより、帯域分割部の演算処理を約 $q/8$ だけ軽減することが可能となる。従って、CPUの性能不足などにより、従来の方法では、音声入力に伴った実時間符号化処理が困難または不可能な場合にも、演算処理判定定数 q の設定により負担を軽減することで実時間処理を行うことが可能となる。なお、演算処理判定定数 q の値の決定については、実施の形態9における単位期間判定定数の決定と同様に行うことができる。

【0320】実施の形態11. 本発明の実施の形態11による音声符号化装置は、帯域分割した信号の一部につ

いて後段の処理を省略することで、処理負担の軽減を図るものである。図26は、本発明の実施の形態11による音声符号化装置の構成を示すブロック図である。図26に示すように、本実施の形態11による音声符号化装置は、音声入力部1101、入力音声サンプリング部1103、帯域分割部1105、符号化ビット割り当て部1106、量子化部1107、符号化部1108、符号化データ記録部1109、レジスタ1112、および帯域間引き部1118から構成されている。また、本実施の形態11の装置も実施の形態9と同様、図11に示されるハードウェア構成である。

【0321】同図において、帯域選択定数レジスタ1112は、メインメモリまたは外部記憶装置で実現され、帯域選択定数を記憶する。帯域間引き部1118は、CPU、メインメモリ、およびプログラムで実現され、帯域分割部1105が分割した帯域信号より、レジスタ1112に記憶された帯域選択定数に基づいて、選択抽出を行う。音声入力部1101、入力音声サンプリング部1103、帯域分割部1105、符号化ビット割り当て部1106、量子化部1107、符号化部1108、および符号化データ記録部1109は、実施の形態9の901、903、および905～909と同様であり、説明を省略する。

【0322】図27は本実施の形態11による音声符号化のフローチャート図、図28は本実施の形態11による音声符号化を説明するための概念図である。以下に本実施の形態11による音声符号化の際の動作を、図27のフローチャートに従って、図26、および図28を参照しながら説明する。

【0323】図27のフローのステップ1～2は実施の形態9と同様に実行され、音声入力部1101より入力された音声信号が、入力音声サンプリング部1103においてサンプリングされ、得られたサンプリングデータに対して、帯域分割部1105がM個の周波数帯域への分割を行い、M個の帯域信号データが得られる。

【0324】次のステップ3において、帯域間引き部1118は、レジスタ1112より、記憶された帯域選択定数rを得て、帯域分割部1105の出力するM個の帯域信号データに対してr個おきに帯域信号データを選択取得し、計 $M/(r+1)$ 個の帯域信号データを抽出する。rは0以上の整数として、装置性能等に応じて予め設定され、レジスタ1112に記憶されたものである。ここで $r=2$ とした場合には、図28に示すように2個おきに○印をつけた帯域信号データが抽出される。帯域間引き部1118は抽出した $M/(r+1)$ 個の帯域信号データを量子化部1107に出力する。

【0325】ステップ4において、符号化ビット割り当て部1106は、レジスタ1112より取得した帯域判定定数rに基づいて、ステップ3で抽出された帯域のみに対して符号化ビット割り当て数を決定する。決定され

た $M/(r+1)$ 個の帯域に対しての符号化ビット割り当て数は、符号化ビット割り当て部1106から量子化部1107に伝えられる。ステップ5以降は $M/(r+1)$ 個のデータに対して第1の従来例と同様に実行される。

【0326】このように、本実施の形態11による音声符号化装置においては、帯域選択定数rを記憶するレジスタ1112と、帯域間引き部1118とを備え、帯域間引き部1118は、帯域選択定数rに基づき、帯域分割処理で得られたM個の帯域信号データより、 $M/(r+1)$ 個の帯域信号データを抽出し、後段の処理はこの抽出されたデータに対して行われるので、帯域選択定数rの値を制御することにより、符号化ビット割り当て、および量子化の処理を約 $1/r$ に軽減することが可能となる。但し、rが1以上の場合がこれに該当するものであり、 $r=0$ とした場合では処理負担は変わらない。

【0327】従って、CPUの性能不足などにより、従来の方法では、音声入力に伴った実時間符号化処理が困難または不可能な場合にも、帯域選択定数rの設定により負担を軽減することで実時間処理を行うことが可能となる。なお、帯域選択定数rの値の決定については、実施の形態9における単位期間判定定数の決定と同様に行うことができる。

【0328】実施の形態12。本発明の実施の形態12による音声符号化装置は、入力データ量を監視し、制御用の定数をそれに応じて変更し得るものである。図29は、本発明の実施の形態12による音声符号化装置の構成を示すブロック図である。図29に示すように、本実施の形態9による音声符号化装置は、音声入力部1201、レジスタ1202、入力音声サンプリング部1203、判定制御部（単位期間判定）1204、帯域分割部1205、符号化ビット割り当て部1206、量子化部1207、符号化部1208、符号化データ記録部1209、固定的符号レジスタ1210、入力バッファ1213、固定的符号レジスタ1210、および入力バッファ監視部1214から構成されている。すなわち、実施の形態9に入力バッファ1213と入力バッファ監視部1214とを追加した構成となっている。また、本実施の形態12の装置も実施の形態9と同様、図11に示されるハードウェア構成である。

【0329】同図において、入力バッファ1213は、メインメモリ等のメモリで実現され、データを一時記憶する。入力バッファ監視部1214は、CPU、メインメモリ、およびプログラムで実現され、入力バッファ1213に一時記憶のため保持されるデータ量を調べて、このデータ量を予め設定された値と比較し、その結果によって、レジスタ1202の定数kの値を変更する。レジスタ1202は、記憶する定数の値が入力バッファ監視部1214によって変更されることを除いて実施の形態9のレジスタ902と同様である。入力音声サンプリ

ング部1203は、サンプリング音声データを入力バッファ1213に出力する点を除いて、実施の形態9の入力音声サンプリング部903と同様である。判定制御部1204は、入力バッファ1213からサンプリング音声データを取りだして処理対象とする点を除いて、実施の形態9の判定制御部904と同様である。また、音声入力部1201、帯域分割部1205、符号化ビット割り当て部1206、量子化部1207、符号化部1208、符号化データ記録部1209、および固定的符号レジスタ1210は実施の形態9における901、および905~910と同様である。

【0330】図30は本実施の形態12による音声符号化のフローチャート図である。以下に本実施の形態12による音声符号化の際の動作を、図30のフローチャートに従って、図29を参照しながら説明する。ここで、単位期間判定定数 k としては、CPUの性能に基づいて予め定められた値「2」が初期値としてレジスタ2に格納されているものとする。

【0331】図30のフローのステップ1で、音声入力部1201より入力された音声信号は、入力音声サンプリング部1203において、実施の形態9と同様にサンプリングされ、ステップ2において、サンプリングデータは入力バッファ1213に書き込まれて一時記憶される。ステップ3において、判定制御部1204は入力バッファ1213より、一時記憶されたサンプリングデータを読み出す。

【0332】続いて行われるステップ4については後述する。ステップ5の判定は、判定制御部1204がステップ3で読み出したデータに対して実施の形態9の場合と同様に行われ、これ以降、ステップ9の符号化データ出力までとステップ10は、実施の形態9における図41のフローのステップ3~6およびステップ7と同様に実行されるので、ステップ5~9およびステップ10については説明を省略する。これらステップの実行により、ここでは2となっている単位期間判定定数 k の数値の設定に応じて、単位期間 t_1 、 t_3 …は符号化対象として帯域分割、符号化ビット割り当て、量子化の処理を行い、 t_2 、 t_4 …には上記処理を行わず、処理負担を軽減した符号化が実行される。

【0333】本実施の形態12においては、ステップ3が実行された後のステップ4で、入力バッファ監視部1214が、入力バッファ1213のデータ量を調べて、このデータ量を予め設定した値と比較し、比較結果に基づいて、レジスタ1202に記憶される定数 k の値を変更する。かかるデータ量に基づく定数 k の値の制御には様々な方法を採用し得るが、ここでは以下の様に行われるものとする。

【0334】CPUの負担増大などにより、当初の設定では、音声入力に伴っての符号化処理が出来なくなった場合、入力バッファ1213については、書き込みは同

じペースで行われるのに対して、符号化処理が遅滞するのに伴ってそのための読み出しのペースが落ちるので、蓄積されるデータ量は増大する。

【0335】従って、入力バッファ監視部1214は、入力バッファ1213のデータ量が、予め設定した値であるバッファフルレベルBFを越えた場合は、現在の設定での実時間符号化処理が不可能であると判断し、レジスタ1202に記憶される定数 k の値を、1だけ増加させて $k=3$ に変更する。それ以後のフロー図のステップ5~9においては、期間 t_2 と t_3 に関しては、帯域分割、符号化ビット割り当て、量子化の処理を行わず、帯域出力0として、予め設定された符号化データ d_N を出力する。最初の設定である $k=2$ の時は、2回に1回の割合で、帯域分割、符号化ビット割り当て、量子化の処理を省くことにより、これらの処理の負担を約 $1/2$ に軽減できていたが、 $k=3$ に増やすことにより、3回に2回の割合で、帯域分割、符号化ビット割り当て、量子化の処理を省くことにより、これらの処理の負担を約 $1/3$ と、さらに軽減させることができる。このようにして、入力バッファ監視部1214は定数 k の値を変更することにより、CPUに対する負担の軽減を図り、音質の低下は伴うものの実時間処理を継続することが可能となる。

【0336】図30のフローにおける繰り返しで、ステップ4において、なおも入力バッファ1213のデータ量がバッファフルレベルBFを越える場合、入力バッファ監視部1214は、レジスタ1202の定数 k を変更し、さらに1増加して $k=4$ とする。これにより、ステップ5~9においては、期間 t_2 から t_4 に関しては、帯域分割、符号化ビット割り当て、量子化の処理を行わず、帯域出力0として、予め設定された符号化データ d_N を出力することにより、帯域分割、符号化ビット割り当て、量子化の処理の負担を約 $1/4$ に軽減させる。以後、ステップ4での入力バッファのデータ量がバッファフルレベルBF以下になるまで、入力バッファ監視部1214は、レジスタ1202の k の値を増加させる。

【0337】逆にCPUの負担が減少するなど、音声入力に伴っての符号化処理を行ってもなお、CPUの処理能力に余力が生じた場合、入力バッファ1213からの読み出し量が多くなるので、蓄積されるデータ量は減少し、ついには、少量のデータが短時間蓄積されるのみの状態が続くに至ることとなる。

【0338】従って、ステップ4において、入力バッファ1213のデータ量が予め設定した値であるバッファエンptyレベルBEを下回る場合は、入力バッファ監視部1214は、符号化処理能力に余力があると判断する。なるべく定数 k の値が少ない方が、帯域出力0の時間が短く、高品質の符号化データが得られるので、入力バッファ監視部1214は、定数 k の値を1だけ減少させ、以後は上記と同様に入力バッファ1213のデータ

量がバッファエンブティレベル B E 以上になるまで、図 30 のフローの繰り返しにおいて、ステップ 4 でレジスタ 1202 の定数 k の値を 1 ずつ減少させる。

【0339】なお、上記の方法では、定数 k の値を制御するために、バッファフルレベル B F とバッファエンブティレベル B E の 2 つの値を用いたが、バッファフルレベル B F のみを用いてもよく、この場合、例えば入力バッファのデータ量が予め設定したバッファフルレベル B F に達するまで定数 k の値を増加し、音声入力と符号化処理がつりあうとき、すなわちデータ量が B F に達したときに、定数 k を増加するのを止めるように制御を行う等のように制御を実現できる。

【0340】このように、本実施の形態 12 の音声符号化装置においては、実施の形態 9 の音声符号化装置に入力バッファ 1213 と、入力バッファ監視部 1214 とをさらに備えたものとしたことで、サンプリングデータをこの入力バッファ 1213 に一時記憶した後に読み出して、それ以後の処理を行うものとし、また、入力バッファ監視部 1214 が、入力バッファ 1213 のデータ量を調べることにより、レジスタに記憶する単位期間判定定数 k の値を動的に制御することによって、CPU の基本的性能に応じた実時間符号化処理を行えることに加え、CPU の処理性能の変化にも対応して、その時点で符号化処理可能な、最も高品質な音声符号化を行うように図ることが可能となる。

【0341】したがって、音声符号化装置として汎用のパーソナルコンピュータ等を用いて、マルチタスク下で音声符号化処理を実行するような場合においても、他のタスク実行による CPU の処理能力の変化に対応して、実時間で符号化処理を実行することができる。

【0342】なお、ここでは、音声符号化装置の構成を、実施の形態 9 による装置に入力バッファ 1213 と、入力バッファ監視部 1214 とを追加する構成としたものであるが、実施の形態 10 による装置に対して入力バッファ 1213 と、入力バッファ監視部 1214 とを追加する構成とすることも可能であり、演算処理判定定数 q の値を制御することにより、基本低域通過フィルタ処理の演算量を増減して、処理負担軽減や、符号化データの音質向上を図ることが可能である。

【0343】同様に、実施の形態 11 による装置に対して入力バッファ 1213 と、入力バッファ監視部 1214 とを追加する構成とすることも可能であり、帯域選択定数 r の値を制御することにより、選択抽出される帯域信号データの数を増減して、処理負担軽減や、符号化データの音質向上を図ることが可能である。

【0344】実施の形態 13. 本発明の実施の形態 13 による音声符号化装置は、出力する符号化データ量を監視し、制御用の定数をそれに応じて変更し得るものである。図 31 は、本発明の実施の形態 13 による音声符号化装置の構成を示すブロック図である。同図に示すよう

に、当該音声符号化装置は、音声入力部 1301、レジスタ 1302、入力音声サンプリング部 1303、判定制御部 1304、帯域分割部 1305、符号化ビット割り当て部 1306、量子化部 1307、符号化部 1308、符号化データ記録部 1309、固定的符号レジスタ 1310、および符号化データ監視部 1315 から構成されている。この構成は、実施の形態 9 による音声符号化装置に符号化データ監視部 1315 を追加した構成である。また、本実施の形態 13 の装置も実施の形態 9 と同様、図 11 に示されるハードウェア構成である。

【0345】符号化データ監視部 1315 は、CPU、メインメモリ、およびプログラムで実現され、符号化部 1308 より出力される単位時間当たりの符号化データ量を調べて、このデータ量を予め設定された値と比較し、その結果によって、レジスタ 1302 の定数 k の値を変更する。レジスタ 1302 は、記憶する定数の値が符号化データ監視部 1315 によって変更されることを除いて実施の形態 9 のレジスタ 902 と同様である。音声入力部 1301、入力音声サンプリング部 1303、判定制御部 1304、帯域分割部 1305、符号化ビット割り当て部 1306、量子化部 1307、符号化部 1308、符号化データ記録部 1309、固定的符号レジスタ 1310 は実施の形態 5 における 901、および 903~910 と同様である。すなわち、実施の形態 12 の装置が入力データの量を監視したのに対して、本実施の形態 13 の装置では符号化データの量を監視して、レジスタに記憶された定数の値を制御するものである。

【0346】図 32 は本実施の形態 13 による音声符号化のフローチャート図である。以下に本実施の形態 13 による音声符号化の際の動作を、図 32 のフローチャートに従って、図 31 を参照しながら説明する。ここで、単位期間判定定数 k としては、CPU の性能に基づいて予め定められた値「2」が初期値としてレジスタ 2 に格納されているものとする。

【0347】図 32 のフローのステップ 1 からステップ 7 までは、実施の形態 9 と同様に実行されるので、説明を省略する。これらステップの実行により、ここでは 2 となっている単位期間判定定数 k の数値の設定に応じて、単位期間 t1、t3... は符号化対象として帯域分割、符号化ビット割り当て、量子化の処理を行い、t2、t4... には上記処理を行わず、処理負担を軽減した符号化が実行される。

【0348】本実施の形態 13 では、ステップ 1 に戻って繰り返しを行う前にステップ 8 が実行され、符号化データ監視部 1315 が、符号化部 1308 より出力される単位時間当たりの符号化データ量を調べて、このデータ量を予め設定された値と比較し、その結果によって、レジスタ 1302 の定数 k の値を変更する。符号化データ量を監視して、定数 k の値を制御するには様々な方法を採用し得るが、ここでは以下の方法に従って行われる

ものとする。

【0349】CPUの負担増大などにより、当初の設定では、符号化処理が間に合わなくなった場合、符号化処理のペースが落ちるため、出力される符号化データ量は減少する。従って、ステップ8において、符号化データ量が、予め設定した値である符号化最低レベルCLに達しない場合は、符号化データ監視部1315は、実施の形態12に示した入力バッファ監視部1214と同様、レジスタ2に記憶される単位期間判定定数kの値を増加させることにより、CPUの負担を軽減させるように図る。

【0350】同様に、CPUの負担が減少し、余力が生じた場合には、ある限度までは符号化データ量が増加することとなるので、ステップ8において、単位時間当たりの符号化処理量が符号化最高レベルCHを下回らない場合は、高品質の符号化が行えるよう、レジスタ1302の単位期間判定定数kの値を減少させる。

【0351】図32のフローの繰り返しにおいては、実施の形態12と同様に上記のような単位期間判定手数の増減が繰り返し行われ、定数kが適切な値となるように制御がなされる。また、符号化最低レベルCLと符号化最高レベルCHの2つの値を用いず、符号化最低レベルCLのみを用いても制御可能である点については、実施の形態12と同様である。

【0352】このように、本実施の形態13の音声符号化装置においては、実施の形態9による装置に、符号化データ監視部1315をさらに備えたものとしたことで、符号化データ監視部1315が、単位時間当たり出力される符号化データ量を調べることにより、これをその時点におけるCPUの符号化処理能力の指標として、状況に応じてレジスタ1302に記憶する定数kの値を動的に制御することによって、CPUがその時点で、符号化処理可能な、最も高品質な音声符号化を行うように図ることが可能となる。従って、実施の形態12の装置と同様に、マルチタスク等による、その時点でのCPUの処理能力の変化に対応することが可能となる。

【0353】また、本実施の形態13についても、実施の形態12と同様に、実施の形態10による装置、または実施の形態11による装置に、符号化データ監視部1315を追加する構成とすることも可能であり、演算処理判定定数、または帯域選択定数の値を制御することによって、同様の効果が得られる。

【0354】実施の形態14. 本発明の実施の形態14による音声符号化装置は、符号化ビットの割り当て方によって、心理聴覚分析代替制御を実現し、処理負担を大きく増すことなく、符号化データの再生音質の向上を図るものである。図33は、本発明の実施の形態14による音声符号化装置の構成を示すブロック図である。同図に示すように、本実施の形態14による音声符号化装置は、音声入力部1401、入力音声サンプリング部14

03、帯域分割部1405、符号化ビット割り当て部1406、量子化部1407、符号化部1408、符号化データ記録部1409、およびビット割り当て制御部（順次ビット割当）1416から構成されている。また、本実施の形態14の装置も実施の形態9と同様、図11に示されるハードウェア構成である。

【0355】同図において、ビット割り当て制御部1416は、CPU、メインメモリ、およびプログラムで実現され、帯域分割部1405の分割によって得られるM個の帯域信号データに対して、符号化ビット割り当て部1406が割り当てるビット数を、所定のアルゴリズムに従って算定する。音声入力部1401、入力音声サンプリング部1403、帯域分割部1405、符号化ビット割り当て部1406、量子化部1407、符号化部1408、および符号化データ記録部1409は、実施の形態9の901、903、および905～909と同様であり、説明を省略する。

【0356】図34～36は本実施の形態14による音声符号化の動作を示すフローチャート図である。以下に本実施の形態13による音声符号化の際の動作を、図34～36のフローチャートに従って、図33を参照しながら説明する。図34のフローのステップ1～2は第2の従来例における図65のステップ1～2と同様に実行され、M個の周波数帯域に分割された帯域信号データが得られる。MPEGオーディオの規格に従って、第2の従来例と同様にM=32個であったとする。

【0357】この後、図65に示す第2の従来例では、L=256個に分割された帯域信号データに対して高速フーリエ変換と心理聴覚分析とを行って、符号化ビット割り当て数を決定するが、本実施の形態14では、図34のフローのステップ3において、ビット割り当て制御部16が心理聴覚分析代替制御方式として、順次ビット割り当て方式により算定した結果に基づいて符号化ビット割り当て部1406がM=32個の帯域信号データに対する符号化ビットの割り当てを行う。

【0358】まず割り当てられるべき総ビット数は、MPEGオーディオのレイヤ1ならば256kbps、レイヤ2ならば192kbpsと定められたビットレートから求められる。そして求められた総ビット数を以下に説明する順次ビット割り当て方式により、割り当てを行う。

【0359】図35は順次ビット割り当て方式の手順を示すフローチャートである。ビット割り当て制御部16は、ステップ101とステップ103で帯域0～10に対してビットを割り当て、その後ステップ105とステップ107とで帯域11～22に対してビットを割り当てる。すなわち、帯域0～10と、帯域11～22とに各2回ずつビットを割り当てる。その後ステップ109で帯域0～10に更に1回、ステップ111で帯域11～22に更に1回ビットを割り当てる。それから、ステ

ップ113で帯域23~31に1回ビットを割り当てる。以上の各ステップの実行後にある判定ステップ、ステップ102, 104, 106, 108, 110, 112, 114のいずれかで総ビット数を割り当て終わつたと判断されなければ、ステップ101から上記の割り当てが繰り返され、上記いずれかの判定ステップで総ビット数が割り当てられた際にフローが終了する。

【0360】図36は各帯域へのビット割り当ての手順を示すフローチャートであり、図35のフローのステップ101, 103, 109では図36(a)が、図35のステップ105, 107, 111では図36(b)が、そして、図35のステップ113では図36(c)が実行される。

【0361】図36(a)のフローでは、まずステップ201において変数aを0とし、ステップ202で帯域0に1ビットを割り当てる。ステップ203で総ビット数を割り当て終わっていない場合には、ステップ204で $a=0+1=1$ とされ、ステップ205の判定によりステップ202に戻り、帯域1に1ビットが割り当てられる。これを繰り返すことで、ステップ205で $a=11$ と判定された場合、すなわち帯域0~10に各1ビットずつを割り当て終わつたとき、またはステップ203において総ビット数を割り当て終わつたと判定されたとき、図36(a)のフローは終了する。図36(b)、および(c)のフローも同様である。図36(a)~(c)のフローが終了したときは、図35のフローの元のステップに復帰して、総ビット数割り当てが終わっていた場合には直後の判定ステップにより図35のフローも終了する。そして、図34のフローのステップ3が終了し、同図ステップ4~5は第2の従来例と同様に実行される。

【0362】本実施の形態14の装置における順次ビット割り当て方式は、上記のように帯域0~10と、11~22とにビットを割り当てていく方式である。これは第2の従来例において、図66に示すように、これらの帯域については最小可聴限界が低い、すなわち人間の聴覚で聞き取りやすい帯域であり、帯域分割で32の帯域に分割された周波数帯域の中で、人間の耳に聞こえやすい帯域に対してより大きな重み付けをし、各帯域に対する重み付けの大きい順にビットを割り当てるという考え方である。そのため、大きな音圧を有する帯域であるか、音圧がほとんどない帯域であるかにかかわらず、従って、基本的に入力信号にかかわらず、上記のように順次ビットを割り当ててものである。

【0363】このように、本実施の形態14の音声符号化装置においては、ビット割り当て制御部1416を備えたことで、定められたアルゴリズムにより順次帯域にビットを割り当てるので、処理負担を大きく増大することなく、人間の聴覚特性を活かした符号ビット割り当てが実行でき、再生音質の良好な符号化データが得られる。すなわち本実施の形態14の音声符号化装置は、第

1の従来例による装置にビット割り当て制御部1416を追加した構成により、図35~36に示す単純なアルゴリズムに従った簡単な符号化ビット割り当て方法を32の帯域に対して実行することにより、再生音質の向上を図り得るものであり、256の帯域に対してフーリエ変換と心理聴覚分析とを行う第2の従来例と比較すると、はるかに処理負担は小さく、図21のハードウェア構成で示される汎用パーソナルコンピュータやワークステーション等で行う音声符号化においても実時間処理と音質向上との両立が可能なものである。なお、順次ビット割り当て方式のアルゴリズムについては、図35~36に示したものは一例であり、これに限定されるものではなく、帯域の順番や割り当てビット数を変更しても同様の単純な順次割り当てを実行することは可能であり、同様の効果が得られる。

【0364】実施の形態15. 本発明の実施の形態15による音声符号化装置は、符号化ビットの割り当て方によって、心理聴覚分析代替制御を実現し、帯域ごとの出力レベルを考慮することで、符号化データの再生音質の一層の向上を図るものである。図37は、本発明の実施の形態15による音声符号化装置の構成を示すブロック図である。同図に示すように、本実施の形態15による音声符号化装置は、音声入力部1501、入力音声サンプリング部1503、帯域分割部1505、符号化ビット割り当て部1506、量子化部1507、符号化部1508、符号化データ記録部1509、およびビット割り当て制御部(帯域出力適応ビット割当)1517から構成されている。これは、実施の形態14による装置と同等の構成である。また、本実施の形態15の装置も実施の形態9と同様、図11に示されるハードウェア構成である。

【0365】同図において、ビット割り当て制御部1517は、CPU、メインメモリ、およびプログラムで実現され、帯域分割部1505の分割によって得られるM個の帯域信号データに対して、符号化ビット割り当て部6が割り当てるビット数を、所定のアルゴリズムに従って算定する。このようなビット割り当て制御部1517の機能を除いては、本実施の形態15は実施の形態14と同様の構成である。従って、音声入力部1501、入力音声サンプリング部1503、帯域分割部1505、符号化ビット割り当て部1506、量子化部1507、符号化部1508、および符号化データ記録部1509は、実施の形態9の901、903、および905~909と同様であり、説明を省略する。

【0366】図38は本実施の形態15による音声符号化のフローチャート図である。以下に本実施の形態15による音声符号化の際の動作を、図38のフローチャートに従って、図37を参照しながら説明する。図38のフローのステップ1~2は第2の従来例における図65のステップ1~2と同様に実行され、M個の周波数帯域

に分割された帯域信号データが得られる。MPEGオーディオの規格に従って、第2の従来例と同様にM=32個であったとする。

【0367】この後、実施の形態14では、図35のフローのステップ3において、ビット割り当て制御部1416が順次ビット割り当て方式により算定するが、本実施の形態15ではビット割り当て制御部1517が帯域出力適応ビット割り当て方式によりビット割り当てを決定する。上述のように実施の形態14では、帯域分割で32の帯域に分割された周波数帯域に対して、その帯域の有する音圧を考慮に入れずビット割り当てを行うものであったが、本実施の形態15では、人間の耳に聞こえやすい帯域であるかどうかと、各帯域の有する音圧との2つの要因に基づいて各帯域に対するビット割り当て情*

帯域0~10 帯域11~15 帯域16~18 帯域19~22 帯域23以降

5 : 3 : 2 : 1 : 割り当てなし

(2) 各帯域出力レベル比

次に、各帯域の有する音圧を比率として表した、各帯域ごとの出力レベル比を求める。ここでは次のようになっていたものとする。

帯域9,10 帯域11 その他の帯域

3 : 2 : 1

(3) 帯域出力適応重み付け

※
sb0~8 sb9,10 sb11 sb12~15 sb16~18 sb19~22 sb23~31
5 : 15 : 6 : 3 : 2 : 1 : 0

この結果に基づいて、ビット割り当て制御部17は、各帯域のビット割り当てが上記の重み付け(3)に近づくように、総ビット数を割り振ってビット割り当て数を決定し、符号化ビット割り当て部6が符号化ビット割り当てをする。この後図38のフローのステップ4および5は、第2の従来例と同様に実行される。

【0370】このように、本実施の形態15の音声符号化装置においては、ビット割り当て制御部1517を備えたことで、人間の耳に聞こえやすい帯域であるかどうかと、各帯域の有する音圧との2つの要因に基づいて各帯域にビットを割り当てるので、処理負担を大きく増大することなく、人間の聴覚特性を活かした符号ビット割り当てが実行でき、再生音質の良好な符号化データが得られる。

【0371】すなわち本実施の形態15の音声符号化装置は第1の従来例による装置にビット割り当て制御部1517を追加した構成により、比較的単純な演算処理を用いる簡単な符号化ビット割り当て方法を32の帯域に対して実行することにより、再生音質の向上を図り得るものであり、256の帯域に対してフーリエ変換と心理聴覚分析とを行う第2の従来例と比較すると、はるかに処理負担は小さく、図2のハードウェア構成で示される汎用パーソナルコンピュータやワークステーション等で行う音声符号化においても実時間処理と音質向上との両立が可能なものである。また、本実施の形態15による

* 報を生成するものである。

【0368】まず、レイヤ1なら256kbpsというビットレートから、割り当てるべき総ビット数が求められることは実施の形態14と同様である。ビット割り当て制御部17は、「人間の耳に聞こえやすい帯域への重み付け(1)」と、「各帯域出力レベル比(2)」とに基づいて「帯域出力適応重み付け(3)」を求め、これに応じて上記求めた総ビット数を割り当てる。

【0369】(1) 人間の耳に聞こえやすい帯域への重み付け

まず、「人間の耳に聞こえやすい帯域への重み付け」としては、図65に示される最小可聴限界に基づいて次のように決定する。

※次に、考慮すべき二つの要因である上記項目(1)、および(2)から、項目(1)*項目(2)を計算する。これにより次の結果が得られる。(sb:帯域)

20

音声符号化装置は、実施の形態14による装置と比較して、入力音声の特質を要因として処理する分だけ処理負担が大きくなるが、それだけ再生音質の良好な符号化データが得られる。

【0372】なお、帯域出力適応ビット割り当て方式については、本実施の形態15に示した算定方法は一例であり、これに限定されるものではなく、帯域に対する重みづけや、二つの要因の重みを変更しても同様の単純な演算処理による割り当てを実行することは可能であり、同様の効果が得られる。

【0373】実施の形態16. 本発明の実施の形態16による音声符号化装置は、符号化ビットの割り当て方によって、心理聴覚分析代替制御を実現し、各帯域ごとの出力レベルと、各帯域ごとのビット割り当て数とを考慮することで、符号化データの再生音質の一層の向上を図るものである。

【0374】図39は、本発明の実施の形態16による音声符号化装置の構成を示すブロック図である。同図に示すように、本実施の形態16による音声符号化装置は、音声入力部1601、入力音声サンプリング部1603、帯域分割部1605、符号化ビット割り当て部1606、量子化部1607、符号化部1608、符号化データ記録部1609、およびビット割り当て制御部(改良型帯域出力適応ビット割当)1616から構成されている。これは、実施の形態14による装置と同等の

構成である。また、本実施の形態16の装置も実施の形態9と同様、図11に示されるハードウェア構成である。

【0375】図39において、ビット割り当て制御部1617は、CPU、メインメモリ、およびプログラムで実現され、帯域分割部1605の分割によって得られるM個の帯域信号データに対して、符号化ビット割り当て部1606が割り当てるビット数を、所定のアルゴリズムに従って算定する。このようなビット割り当て制御部1617の機能を除いては、本実施の形態16の装置は、実施の形態14、および15の装置と同様の構成である。従って、音声入力部1601、入力音声サンプリング部1603、帯域分割部1605、符号化ビット割り当て部1606、量子化部1607、符号化部1608、および符号化データ記録部1609は、実施の形態9の901、903、および905～909と同様であり、説明を省略する。

【0376】図40～図41は本実施の形態16による音声符号化のフローチャート図である。以下に本実施の形態16による音声符号化の際の動作を、図40～図41のフローチャートに従って、図39を参照しながら説明する。

【0377】図40のフローのステップ1～2は第2の従来例における図65のステップ1～2と同様に実行され、M個の周波数帯域に分割された帯域信号データが得られる。MPEGオーディオの規格に従って、第2の従来例と同様にM=32個であったとする。

【0378】この後、実施の形態15では、図38のフローのステップ3において、ビット割り当て制御部1517が帯域出力適応ビット割り当て方式により算定するが、本実施の形態16ではビット割り当て制御部1617が改良型帯域出力適応ビット割り当て方式によりビット割り当てを決定する。帯域分割で32の帯域に分割された周波数帯域に対して、実施の形態14では、その帯域に割り当てられるビット数を考慮に入れず、人間の耳*

帯域0～10 帯域11～15 帯域16～18 帯域19～22 帯域23以降

1 : 2 : 3 : 5 : 20

(3) 各帯域ごとのビット割り当て数に対応した重み付け次に、「各帯域ごとのビット割り当て数に対応した重み付け」としては、図66に示す最小可聴限界に基づき、さらに同一の帯域に必要な以上のビット割り当てをしない

*に聞こえやすい帯域を優先してビット割り当てを行うものであり、また、実施の形態15では、人間の耳に聞こえやすい帯域であるかどうかと、各帯域の有する音圧との2つの要因に基づいて各帯域に対するビット割り当てを行うものであった。そして、本実施の形態16では、さらに実施の形態15における2つの要因に加えて、各帯域へのビット割り当て数が十分であるかどうかという要因を考慮し、3つの要因に基づいて各帯域に対するビット割り当て情報を生成するものである。

10 【0379】以下に本実施の形態16で用いるビット割当の方法について説明する。まず、レイヤ1なら256kbpsというビットレートから、割り当てるべき総ビット数が求められることは実施の形態14、および15と同様である。

【0380】ビット割り当て制御部1617は、「各帯域出力レベル(1)」と、「人間の耳に聞こえやすい帯域への重み付け(2)」と、「各帯域ごとのビット割り当て数に対応した重み付け(3)」とに基づいて上記求めた総ビット数を割り当てる。

20 【0381】(1) 各帯域出力レベル

まず、各帯域の有する音圧からスケールファクタを求める。なお、スケールファクタは、0から62の間の値を取り、値の小さいものほど音圧は大きい。ここでは、次のようになっていたものとする。

帯域3,4 帯域5,6 その他の帯域

3 : 9 : 37

(2) 人間の耳に聞こえやすい帯域への重み付け

次に、「人間の耳に聞こえやすい帯域への重み付け」としては、図66に示される最小可聴限界に基づいて次のように決定する。なお、上記のように各帯域出力レベルをスケールファクタとして表しているため、実施の形態14で示した重み付け値とは、意味が逆転しているものとなる。

ように考慮して作成した下記の表に従って決定する。

【0382】

40 【表14】

帯域別ビット割当重み付け表

	ビット 割当 数															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
sb 0	1	2	3	4	5	5	6	7	8	10	12	20	99	99	99	99
1	1	2	3	4	5	5	6	7	8	10	12	20	99	99	99	99
2	1	2	3	4	5	5	6	7	8	10	12	20	99	99	99	99
3	3	3	3	3	3	4	4	6	6	8	8	8	12	20	99	99
4	3	3	3	3	3	4	4	6	6	8	8	8	12	20	99	99
5	3	3	3	3	3	4	4	6	6	8	8	8	12	20	99	99
6	3	3	3	3	3	4	4	6	6	8	8	8	12	20	99	99
7	3	3	3	3	3	4	4	6	6	8	8	8	12	20	99	99
8	3	3	3	3	3	4	4	6	6	8	8	8	12	20	99	99
9	3	3	3	3	3	4	4	6	6	8	8	8	12	20	99	99
10	3	3	3	3	3	4	4	6	6	8	8	8	12	20	99	99
11	3	3	4	4	5	12	99	99	99	99	99	99	99	99	99	99
12	3	3	4	4	5	12	99	99	99	99	99	99	99	99	99	99
13	3	3	4	4	5	12	99	99	99	99	99	99	99	99	99	99
14	3	3	4	4	5	12	99	99	99	99	99	99	99	99	99	99
15	3	3	4	4	5	12	99	99	99	99	99	99	99	99	99	99
16	3	3	4	4	5	12	99	99	99	99	99	99	99	99	99	99
17	3	3	4	4	5	12	99	99	99	99	99	99	99	99	99	99
18	3	3	4	4	5	12	99	99	99	99	99	99	99	99	99	99
19	3	3	4	4	5	12	99	99	99	99	99	99	99	99	99	99
20	3	3	4	4	5	12	99	99	99	99	99	99	99	99	99	99
21	3	3	4	4	5	12	99	99	99	99	99	99	99	99	99	99
22	3	3	4	4	5	12	99	99	99	99	99	99	99	99	99	99
23	3	3	4	99	99	99	99	99	99	99	99	99	99	99	99	99
24	3	3	4	99	99	99	99	99	99	99	99	99	99	99	99	99
25	3	3	4	99	99	99	99	99	99	99	99	99	99	99	99	99
26	3	3	4	99	99	99	99	99	99	99	99	99	99	99	99	99
27	3	3	4	99	99	99	99	99	99	99	99	99	99	99	99	99
28	3	3	4	99	99	99	99	99	99	99	99	99	99	99	99	99
29	3	3	4	99	99	99	99	99	99	99	99	99	99	99	99	99
30	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99
31	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99

【0383】以上の(1) から(3) の要因があることを前提に、次に、考慮すべき三つの要因である上記項目(1)

* を計算する。これにより次のビット割り当て情報係数が得られる。(sb: 帯域)

、(2)、および(3) のうち、初めに項目(1) + 項目(2) * 30

sb0 ~2 sb3,4 sb5,6 sb7 ~10 sb11~15 sb16~18 sb19~22 sb23以

降

1 : 4 : 10 : 37 : 39 : 40 : 42 : 57

ビット割り当て制御部1617は、この結果の中から最小の値を持つ帯域を検出し、その帯域に符号化ビット割り当て部6が符号化ビット割り当てを1ビット行う。なお、最小の値を持つ帯域が複数個ある場合には低い帯域を優先して行う。その後、ビット割り当て制御部161※

※7は、ビット割り当てが行われた帯域に対し、項目(3)からビット数に対応した重み付け(+項目(3))を行い、次のような結果を得る。

【0384】

sb0 sb1,2 sb3,4 sb5,6 sb7 ~10 sb11~15 sb16~18 sb19~22 sb23 以降

2 : 1 : 4 : 10 : 37 : 39 : 40 : 42 : 57

ビット割当部1617による、以上の動作は割り当て可能な総ビット数がなくなるまで繰り返われ、符号化ビット割り当てがなされるものであるが、以下にこのようにビット割当を行う、ビット割当部1617による、図40のフローのステップ3での動作について、図41を用いて説明する。

【0385】図41は改良型帯域出力適応ビット割り当て方式の手順を示すフローチャートである。ステップ101でビット割当部1617は、上記のように各帯域のスケールファクタを算出する。これは、上記(1)に相当

する処理である。次のステップ102では、最小可聴限界に基づいた重み付け処理によって、各帯域へのビット割り当て情報係数を算出する。上記(2)に相当する処理である。

【0386】それから、ステップ103で最小のビット割り当て情報係数を持つ帯域を検出し、ステップ104でその帯域に符号化ビットを1つ割り当てる。すなわち、上記の項目(1) + (2) がなされたこととなる。

【0387】次のステップ105において、上記の項目(3)に相当する重み付けがなされ、ステップ104でビ

ット数を割り当てた帯域に対して、その帯域に現在割り当てられているビット数に対応して、(表14)から得られる重み付け係数を加算する。ステップ103~105の動作はステップ106において総ビット数の割り当てが終了したと判断されるまで繰り返され、終了したと判断されたときは図41のフローが終了する。そして図40のフローのステップ3が終了し、同図ステップ4および5は、第2の従来例と同様に実行される。

【0388】このように、本実施の形態16の音声符号化装置においては、ビット割り当て制御部1617を備えたことで、人間の耳に聞こえやすい帯域であるかどうかと、各帯域の有する音圧と、同帯域への必要以上のビット割り当てを避けることとの3つの要因に基づいて各帯域にビットを割り当てるので、処理負担を大きく増大することなく、人間の聴覚特性を活かした符号ビット割り当てが実行でき、再生音質の良好な符号化データが得られる。

【0389】すなわち本実施の形態16の音声符号化装置は第1の従来例による装置にビット割り当て制御部1617を追加した構成により、比較的単純な演算処理を用いる簡単な符号化ビット割り当て方法を32の帯域に対して実行することにより、再生音質の向上を図り得るものであり、256の帯域に対してフーリエ変換と心理聴覚分析とを行う第2の従来例と比較すると、はるかに処理負担は小さく、第2図のハードウェア構成で示される汎用パーソナルコンピュータやワークステーション等で行う音声符号化においても実時間処理と音質向上との両立が可能なものである。また、本実施の形態16による音声符号化装置は、実施の形態15による装置と比較して、各帯域のビット割り当て状況の監視を行う重み付けを考慮して処理する分だけ処理負担が大きくなるが、それだけ再生音質の良好な符号化データが得られる。

【0390】なお、改良型帯域出力適応ビット割り当て方式については、本実施の形態16に示した算定方法は一例であり、これに限定されるものではなく、帯域に対する重みづけや、各帯域のビット割り当て数に対する重みづけを変更しても、さらには、スケールファクタ値を用いず各帯域出力レベル比を用いても同様の単純な演算処理による割り当てを実行することは可能であり、同様の効果が得られる。

【0391】実施の形態17. 本発明の実施の形態17による音声符号化装置は、最小可聴限界を考慮した符号化ビットの割り当て方によって、心理聴覚分析代替制御を実現し、符号化データの再生音質の向上を図るものである。図42は、本発明の実施の形態17による音声符号化装置の構成を示すブロック図である。同図に示すように、本実施の形態17による音声符号化装置は、音声入力部1701、入力音声サンプリング部1703、帯域分割部1705、符号化ビット割り当て部1706、量子化部1707、符号化部1708、符号化データ記

録部1709、ビット割り当て制御部(動的ビット割り当て)1717、および最小可聴限界比較部1718から構成されている。これは、実施の形態14による装置と同等の構成である。また、本実施の形態17の装置も実施の形態9と同様、図11に示されるハードウェア構成である。

【0392】同図において、最小可聴限界比較部1718は、CPU、メインメモリ、およびプログラムで実現され、帯域分割部1705の分割によって得られるM個の帯域信号データに対して、最小可聴限界との比較を行い、最小可聴限界未満の帯域を検出する。符号化ビット割り当て部1706は、最小可聴限界比較部1718が検出した帯域に対しては符号化ビットを割り当てない。このようなビット割り当て制御部1517の機能を除いては、本実施の形態15は実施の形態14と同様の構成である。従って、音声入力部1701、入力音声サンプリング部1703、帯域分割部1705、符号化ビット割り当て部1706、量子化部1707、符号化部1708、および符号化データ記録部1709は、実施の形態9の901、903、および905~909と同様であり、説明を省略する。

【0393】図43は本実施の形態17による音声符号化のフローチャート図である。以下に本実施の形態17による音声符号化の際の動作を、図43のフローチャートに従って、図42を参照しながら説明する。本実施の形態17による装置では、音声符号化に先立って最小可聴限界比較部1718はメモリ等を内部記憶手段として用いて、M帯域(ここでは32帯域)に対する最小可聴限界をテーブル(表)として記憶しておくものである。このテーブルについては、図66に示すグラフや、このようなグラフを数表化したものより読みとって記憶しておくこととする。

【0394】図43のフローのステップ1~2は第2の従来例における図65のステップ1~2と同様に実行され、M個の周波数帯域に分割された帯域信号データが得られる。MPEGオーディオの規格に従って、第2の従来例と同様にM=32個であったとする。

【0395】この後、図65に示す第2の従来例では、L=256個に分割された帯域信号データに対して高速フーリエ変換と、最小可聴限界との比較を含む心理聴覚分析とを行って、符号化ビット割り当て数を決定するが、本実施の形態17では、図43のフローのステップ3において、最小可聴限界比較部1718が、M=32個の帯域に対して最小可聴限界との比較をし、その比較の結果に基づいて符号化ビット割り当て部1706がM=32個の帯域信号データに対する符号化ビットの割り当てを行う。

【0396】そして、ステップ3での比較では、帯域分割部1705が分割によって得たM=32帯域の帯域信号データについて、最小可聴限界比較部1718が、上

述のように予め記憶したテーブルの対応する帯域の最小可聴限界との比較を行い、最小可聴限界に満たない帯域を抽出して、その結果を符号化ビット割り当て部1706に出力する。

【0397】そして、ステップ4において符号化ビット割り当て部1706は上記出力された比較の結果を用いて、最小可聴限界未満の帯域にはビット割り当てを行わず、その分を最小可聴限界以上の他の帯域に多くのビットを割り当てるように、ビット割り当てを実行する。同図ステップ4および5は、第2の従来例と同様に実行される。

【0398】このように、本実施の形態17の音声符号化装置においては、最小可聴限界比較部1718を備え、帯域分割部1705が分割して得られたM個の帯域に対して、予め記憶した最小可聴限界との比較を行うことで、最小可聴限界未満の帯域を検出し、符号化ビット割り当て部1706は、上記検出した帯域に符号化ビットを割り当てないので、処理負担を大きく増大することなく、人間の聴覚特性を活かした符号ビット割り当てが実行でき、再生音質の良好な符号化データが得られる。

【0399】すなわち本実施の形態17の音声符号化装置は第1の従来例による装置に最小可聴限界比較部1718を追加した構成により、先に分割して得られたM=32個の帯域信号と最小可聴限界との比較を行うものなので、MPEG1Audioにおいて第2の従来例における最小可聴限界適用との比較では、256帯域へのFFTが不要となり、また帯域信号の演算や比較が $32/256=1/8$ に削減でき、大幅な処理負担軽減を図ることが可能となる。従って、図2のハードウェア構成で示される汎用パーソナルコンピュータやワークステーション等で行う音声符号化においても実時間処理と音質向上との両立が可能なるものである。

【0400】なお、音声符号化に先立って最小可聴限界テーブルの記憶については、図66のグラフまたは数表を読み込ませるものとしたが、他に、規格書(ISO/IEC1172-3)の表D.1に従って、各帯域の最小可聴限界を求め、これを記憶しておくこととしても良い。この表では、INDEXと最小可聴限界とが対照されたものとなっているので、M=32帯域であれば、32帯域のそれぞれの中心周波数に近いINDEXの値を用いて、表より最小可聴限界を求めることができる。

【0401】以上本発明の実施の形態として、実施の形態9~17を示したが、実施の形態9~13による符号化においては、実質的にオーディオデータの帯域信号データレベルでの間引きや、フィルタ特性の低下を行うため、それに伴い音質が劣化することにはなる。しかしその場合でも、性能の低いCPUによっても、ハードウェア的追加等を要せずソフトウェア的に、MPEGAudioなどの帯域分割符号化データを実時間で作成でき、これを、動画符号化の国際標準として広く用いられるM

PEGデータとして利用することが可能となる。また、変数定数の値を調整することで、CPUの符号化処理性能にあわせて、間引き具合やフィルタ特性を制御できるため、高性能なCPUのみならず性能が不十分なCPUでもその符号化処理能力なりの音質で符号化することが出来、幅広い性能レベルのCPUで符号化処理が実現できる。

【0402】また、実施の形態14~17による符号化においては、第2の従来例における心理聴覚分析を行う場合ほどの音質向上の効果は得られない。しかし、心理聴覚分析を全く行わない第1の従来例による音声符号化よりは音質を向上でき、汎用パーソナルコンピュータやワークステーション等の機器においても、ハードウェア的追加等を要せずソフトウェア的に、音声取り込みにもなった実時間符号化を実行しつつ、再生音質の向上を図ることが可能である。

【0403】但し、実施の形態9~17のいずれにおいても、ハードウェア面に関しては、CPUが高性能であるほど、またサウンドボードの機能や装置内でのデータ伝送速度が高いほど、高品質な符号化が可能である。また、実施の形態9~17の音声符号化は、音声符号化制御プログラムとして記録媒体に記録し、パーソナルコンピュータ、ワークステーションその他の装置において実行することが可能である。

【0404】また、実施の形態9~17では、符号化データを記憶装置に保存することとしたが、ネットワーク等を介して他の機器に伝達し、他の機器において記録または利用することも可能である。また、実施の形態9~17では、CPU制御によるソフトウェア処理で実現するものとして説明したが、CPUの代わりにDSPが制御するソフトウェア処理によっても、同様である。

【0405】実施の形態18、本発明の実施の形態18による映像音声符号化装置は、汎用計算機等において映像音声の符号化処理をソフトウェア処理によって行う場合に、当該計算機等において負担増大があった場合にも、音声の途切れを防ぐことを、音声データの蓄積量を指標として、映像情報の符号化処理停止を行うものである。

【0406】図44は本発明の実施の形態18による映像音声符号化装置の概略構成を示す図である。図示するように、本実施の形態18による映像符号化装置は、ビデオカメラ1801、音声キャプチャ部1802、音声バッファリング部1803、音声符号化部1805、映像キャプチャ部1806、映像符号化部1807、および符号化負荷評価部1808から構成されている。当該映像音声符号化装置からは、図示するように符号化音声情報と符号化映像情報とが装置出力として出力され、これらは必要に応じて伝送されたり記録されることがなる。

【0407】同図において、ビデオカメラ1801は、

映像音声情報を取り込み、アナログ音声情報とアナログ映像情報とに分けて出力する。音声キャプチャ部1802は、ビデオカメラ1801から出力されたアナログ音声情報を入力し、離散的なデジタルデータからなるデジタル原音声情報として出力する。音声バッファリング部1803は、音声キャプチャ部1802から出力されたデジタル原音声情報を一時的に蓄積する。音声バッファリング部1803に蓄積された原音声情報の総量は、原音声バッファ量1804であり、本実施の形態18による映像音声符号化装置において制御に用いられる情報である。音声符号化部1805は、音声バッファリング部1803に蓄積された原音声情報を取り出して、圧縮符号化処理し、符号化音声情報を出力する。音声符号化部1803は、一時蓄積された原音声情報の取り出しにあたっては、該一時蓄積された原音声情報のうち、もっとも先(過去)に蓄積された原音声情報を取り出して、これを音声バッファリング部1803から削除する。従って、音声バッファリング部1803は、FIFO(First In First Out)構造をとることが望ましく、具体的にはリングバッファ等のアルゴリズムで実現される。

【0408】映像キャプチャ部1806は、ビデオカメラ1801から出力されたアナログ映像情報を入力し、離散的なデジタルデータからなり、単位時間ごとの静止画像の複数枚から構成されるデジタルの原映像情報を出力する。ここで、原映像情報は、あらかじめ設定された解像度を有するものとして得られる。映像キャプチャ部1806と、前出の音声キャプチャ部1802とは、マルチメディア対応タイプのパーソナルコンピュータであれば、一般的に装備されるビデオキャプチャボードとして実現される。映像符号化部1807は、映像キャプチャ部1806から出力された原映像情報を入力し、圧縮符号化して符号化映像情報を出力する。符号化負荷評価部1808は、当該映像音声符号化装置の符号化処理における負荷を評価し、その評価に対応して、映像キャプチャ部1806から出力される原映像情報の映像符号化部1807における処理を制御する。符号化負荷評価部1808による制御は、符号化負荷基準情報1810を用いて、符号化負荷評価情報1809を演算によって取得し、当該取得した符号化負荷評価情報1809の値によって、原映像情報を映像符号化部1807に入力するか、原映像情報を破棄するかを選択することで行われる。そして、原映像情報が破棄されたときには、映像符号化部1807の符号化処理は中断されることとなり、当該符号化装置の計算機資源(CPU時間)が音声符号化部1805に明け渡されることとなる。

【0409】符号化負荷評価部1808による、符号化負荷評価情報1809の取得は、原音声バッファ量1804の値を用いて評価情報を計算し、この評価情報に対して符号化負荷基準情報1810を乗算することによって行われる。本実施の形態18では評価情報の計算にあ

たって、原音声バッファ量1804が音声バッファリング部1803における蓄積可能量の半分以上になれば、評価情報を0%とし、半分以下になれば、評価情報を100%とするものである。

【0410】符号化負荷基準情報110は、映像符号化処理の処理量の基準を示す情報であり、例えば「音声バッファが空である場合に映像の処理をどの程度行うかを示す量」として設定しておくものであるが、本実施の形態18においては常に「1」の値であるものとして、評価情報と符号化負荷基準情報1810との乗算にあたって、上記評価情報がそのまま符号化負荷評価情報1809となるようにしている。したがって、符号化負荷評価情報1809は、0%か100%であり、100%の場合、符号化負荷評価部1808は、その時点で入力された原映像情報を100%映像符号化部1807に入力し、0%の場合、原映像情報をすべて破棄する。そして、映像符号化部1807の処理を中断し、計算機資源(CPU時間)を音声符号化部1805に明け渡すように制御を行う。

【0411】このように構成された、本実施の形態18による映像音声符号化装置の動作の概略は、以下のようになる。すなわち、当該映像音声符号化装置において、ビデオカメラ1801から出力されたアナログ音声情報が入力されると、音声キャプチャ部1802は、離散的なデジタルデータからなるデジタル原音声情報として出力する。この音声キャプチャ部1802から出力されたデジタル原音声情報は、音声バッファリング部1803に一時的に蓄積される。そして、音声符号化部1805は、音声バッファリング部1803に蓄積された原音声のうち、もっとも先(過去)に蓄積された原音声を取り出し、取り出した原音声を音声バッファリング部1803から削除し、原音声を圧縮符号化して符号化音声情報として出力する。音声符号化部1805は、音声バッファリング部1803に蓄積された原音声の総量を示す値である原音声バッファ量1804を更新し、原音声バッファ量1804は、当該映像音声符号化装置による符号化処理のための情報として保持される。

【0412】また、ビデオカメラ1801から出力されたアナログ映像情報が入力されると、映像キャプチャ部1806は、離散的なデジタルデータからなり、予め定義された解像度を持つ単位時間ごとの複数の静止画像情報で構成されるデジタル原映像情報として出力する。この映像キャプチャ部1806から出力された原映像情報が入力されると、映像符号化部1807は、圧縮符号化して符号化情報として出力する。そして、符号化負荷評価部1808は、符号化負荷評価情報1809を計算し、この計算された符号化負荷評価情報1809の値に従って、原映像情報を映像符号化部1807に入力するか、原映像情報を破棄して映像符号化部1807の処理を中断し、計算機資源(CPU時間)を音声符号化部1

805に明け渡すか、否かの動作を決定する。映像符号化部1807は、原映像情報を入力されたなら当該原映像情報を圧縮符号化処理し、符号化映像情報を出力する。

【0413】図45は、本実施の形態18による映像音声符号化装置において、ある映像音声をとりこんで符号化する際の動作を図解的に表した図である。ここで、本装置は、パーソナルコンピュータ等の汎用計算機において実現されるものとし、該汎用計算機は複数の作業（タスク）を並行的に実行できるマルチタスクオペレーションシステムによって作動しているものであって、当該映像音声符号化処理は、オペレーションシステム上で、映像符号化、および音声符号化の各タスクとして扱われるものであるとする。映像符号化、および音声符号化を含む各タスクは、オペレーションシステムによって、計算機資源であるCPU時間の割り当てがなされ、該割り当てられた期間に、CPUの制御によってそれぞれの処理を実行することができる。ここでは、各タスクが一連の処理を終了して割り当てられた計算機資源（CPU時間）の解放を行ったとき、オペレーションシステムは他のタスクに割り当てを行うという制御がされるものとする。

【0414】同図においては、上から下に時間が進行していくものとし、図中の四角形は、マルチタスクオペレーションシステム上の各プロセス（タスク）が計算機資源（CPU時間）を消費していることを示す。四角形間を結ぶ破線矢印は、プロセスがスイッチしたことを示す。破線矢印は、斜め線になっているが、この斜め線の角度がプロセススイッチにかかる時間、すなわちマルチタスクオペレーションシステムのタスク切り替えのためのオーバーヘッドを示す。なお、これ以降の説明では、このオーバーヘッドについては、各タスクにおける処理との比較において相対的にわずかなものとして、説明の上で無視するものとする。

【0415】同図中、「映像符号化プロセス」で示される欄は、映像情報の符号化のためのプロセスが消費する時間を示し、上記構成によれば、符号化付加評価部1808の処理と、映像符号化部1807の処理とを実行するプロセスの作業時間を示す。また、「音声符号化プロセス」で示される欄は、音声情報の符号化のためのプロセスが消費する時間を示し、上記構成によれば、音声符号化部1805の処理を実行するプロセスの作業時間を示す。さらに、「その他のプロセス」は、「映像符号化プロセス」および「音声符号化プロセス」以外の、あらゆるプロセスの作業時間を示す。また、「原音声バッファ量」は、その時間での原音声バッファ量1804を、最大バッファ量（音声バッファリング手段1803における蓄積可能量）に対する割合で示したものである。

【0416】なお、当該映像音声符号化装置の構成において、ビデオカメラ1801は、当該装置を実現する汎

用計算機と接続して、比較的独立して機能する周辺機器であり、上記のようなCPU時間の割り当てをうけ、CPU制御により実行されるプロセスと概ね並行して動作がなされるものである。また、音声キャプチャ部1802、および映像キャプチャ部1806を実現するビデオキャプチャボードについても同様に、比較的独立して機能し得るものであり、音声キャプチャ部1802、および映像キャプチャ部1806についても、上記の各プロセスと概ね並行して動作がなされるものとなる。

【0417】すなわち、映像、または音声の符号化プロセスや、他のプロセスが実行されているときにも、概ね並行して映像音声の取り込みと、デジタル化による原映像情報、および原音声情報の作成、そして、原音声情報の音声バッファリング手段1803への蓄積は行われるものである。

【0418】以下に、この例における本実施の形態18による映像音声符号化装置の動作を、図44、および図45を参照しながら説明する。まず、ビデオカメラ1801が映像音声情報を取り込み、アナログ音声情報とアナログ映像情報とに分けて出力する。アナログ映像情報は、映像キャプチャ部1806に入力され、映像キャプチャ部1806は、アナログ/デジタル変換処理によって、上記のように複数の静止画像情報からなるデジタル原映像情報とし、これを出力する。この過程（プロセス）は主としてビデオカメラ1801の動作と、映像キャプチャ部1806であるキャプチャボードで実行される処理が主体となるものなので、図45に示すような、CPU時間を消費するオペレーティングシステム上の各プロセスと概ね並行的に処理される。

【0419】符号化負荷評価部1808は、一旦出力された原映像情報を入力し、その時点の原音声バッファ量1804を確認する。ここでは、未だ音声は音声バッファリング部1803に入力されておらず、原音声バッファ量1804は0%であるとする。従って、予め決められた基準値である50%を下回っており、上述の通り本実施の形態では、符号化負荷基準情報1810については、値を「1」として、乗算では考慮しなくてもさしつかえないものとしているので、符号化負荷評価基準情報1809は評価情報のままの、100%となる。そこで、符号化負荷評価部1808は、入力した原映像情報の100%を映像符号化部1807に出力する。映像符号化部1807では、原映像情報に対して映像符号化処理を行い、終了した時点でCPU時間を解放する。符号化負荷評価部1808と、映像符号化部1807とによる、以上の処理は、図45の映像符号化プロセスであるA部分に相当する。

【0420】一方、音声キャプチャ部1802は、ビデオカメラ1801より出力されたアナログ音声情報を入力し、アナログ/デジタル変換処理によって、デジタル原音声情報として出力する。原音声情報は、音声バッ

アリング手段に入力されて一時蓄積され、音声バッファリング部1803は、蓄積量に応じて、当該映像音声符号化装置の保持する原音声バッファ量1804を更新する。この過程も主としてビデオカメラ1801の動作と、音声キャプチャ部1802であるキャプチャボードで実行される処理が主体となるものなので、図45に示すようなオペレーティングシステム上の各プロセスと概ね並行的に処理される。ここでは、映像符号化プロセスAと並行してこの過程が実行され、原音声バッファ量が30%に達していたものとする。

【0421】音声符号化部1805は、音声バッファリング部1803から一定量（原音声読み出し量）の原音声情報を、先（過去）に蓄積した分から読み出し、当該読み出した分の原音声情報を音声バッファリング部1803から削除して、原音声バッファ量1804を更新する。さらに、音声符号化部1805は、原音声情報を符号化する。本実施の形態18では、上記の原音声読み出し量について、最大バッファ量の30%とするので、上記の通り、音声バッファリング部1803に音声量が30%分蓄積されていたことから、そのすべてを読み出して

符号化し、符号化が終了した時点でCPU時間を解放する。音声符号化部1805によるこの処理は、図45の音声符号化プロセスであるB部分に相当する。

【0422】ここで、図45に示すように、偶然「その他のアプリケーション」が起動し、CPU時間を要求したので、「その他のアプリケーション」がCPU時間を消費する。「その他のアプリケーション」は、比較的処理負担が大きなものであって、しばらくの間CPU時間を占有してから解放する。他の作業の処理にかかるこの過程（プロセス）は、図45のその他のプロセスである

C部分に相当する。C部分のプロセスとも並行して、ビデオカメラ1801と、音声キャプチャ部1802、および映像キャプチャ部1806による処理は実行されているものとする。従って、原音声情報の一時蓄積がなされ、原音声バッファ量は図45に示すように60%に達する。

【0423】次に、再び符号化負荷評価部1808にCPU時間の割当が回ってきたが、この時点での原音声バッファ量1804である60%は、基準値の50%以上となっていたので、符号化負荷評価部1808が取得する評価情報は0%となり、それに符号化負荷基準情報110の「1」を乗算しても、符号化負荷評価情報1809は0%となる。そこで、符号化負荷評価部1808は、この時点の原映像情報を破棄し、映像符号化部1807による符号化処理は行われることなく、CPU時間はすみやかに解放されることとなる。符号化負荷評価部1808のこの処理は、図45のD部分に相当する。

【0424】そして、音声符号化部1805は、CPU時間の割当が回ってきたので、音声バッファリング部1803から30%の原音声情報を読み出し、その分の原

音声情報を音声バッファリング部1803から削除し、原音声バッファ量1804を更新する。原音声バッファ量は60%から30%になる。さらに、音声符号化部1805は原音声情報を符号化し、符号化が終了した時点でCPU時間を解放する。音声符号化部1805によるこの処理は、図45のE部分に相当する。

【0425】符号化負荷評価部1808にCPU時間の割当がなされる。この時点で原音声バッファ量1804は30%であり、基準値の50%を下回ったので、図45のプロセスA部分での場合と同様に、符号化負荷評価情報1809は100%となり、映像符号化部1807において、原映像情報の符号化が行われる。先の映像符号化処理の場合と比較して、原映像情報が複雑であるために、プロセスA部分の場合よりも符号化処理に時間がかかり、比較的CPU時間を多く消費してから映像符号化プロセスは、CPU時間を解放する。符号化負荷評価部1808と、映像符号化部1807とによる、以上の処理は、図45の映像符号化プロセスであるF部分に該当する。このプロセスFと並行して、原音声情報の蓄積がなされ、原音声バッファ量は90%に達した。

【0426】音声符号化部1805にCPU時間の割当が回ってきたので、音声符号化部1805は、音声バッファリング部1803から30%の原音声情報を読み出し、その分の原音声情報を音声バッファリング部1803から削除し、原音声バッファ量1804を更新する。原音声バッファ量1804は90%から60%になる。さらに、音声符号化部1805は、原音声情報を符号化し、符号化を終了した時点でCPU時間を解放する。音声符号化部1805によるこの処理は、図45のG部分に相当する。

【0427】符号化負荷評価部1808にCPU時間の割当がなされ、この時点で原音声バッファ量1804は90%に達しており、基準値の50%以上であるので、上記のプロセスDの場合と同様に、符号化負荷評価情報1809は0%となり、符号化負荷評価部1808は、この時点の原映像情報を破棄し、映像符号化部1807における映像符号化処理は行われず、すみやかにCPU時間が解放される。符号化負荷評価部1808のこの処理は、図45のH部分に相当する。

【0428】音声符号化部1805にCPU時間が割り当てられたので、音声符号化部1805は、音声バッファリング部1803から30%の原音声情報を読み出し、その分の原音声情報を音声バッファリング部1803から削除し、原音声バッファ量1804を更新する。さらに、音声符号化部1805は、原音声情報を符号化し、符号化の終了した時点でCPU時間を解放する。音声符号化部1805によるこの処理は、図45のI部分に相当する。

【0429】図44において、ビデオカメラ1801より映像音声の取り込みが続いている間は、以上のように

映像符号化、および音声符号化のプロセスが実行されることによって、当該取り込みにともなう映像音声符号化が実行される。そして、映像音声の取り込みの終了後、符号化も終了する。

【0430】図46は、本実施の形態18の映像音声符号化装置における、このような符号化の動作をより長期の時間にわたって説明するための図である。同図において、A区間では音声と映像の符号化処理がバランスよく行われており、音声バッファ量は基準値以下の値を保っているが、その後、図示するように「その他のプロセス」がCPU時間を独占したため、これに並行して原音声情報が過剰に蓄積されることとなる。そこで、続くB区間では、蓄積された原音声情報を処理するべく、音声符号化が優先的に処理される。原音声情報が少なくなり、基準値以下になってからは、同図C区間にみられるように、また平常の処理が実行されることとなる。

【0431】このように、本実施の形態18の映像音声符号化装置によれば、音声バッファリング部1803と、符号化負荷評価部1808とを備え、映像についての符号化対象である原映像情報を入力された符号化負荷評価部1808が、映像符号化部180における符号化処理の前に、その時点での音声バッファリング部1803に蓄積された未処理音声情報の量である原音声バッファ量1804を確認し、十分に小さければ、映像符号化を行い、一定量以上未処理音声情報が蓄積されていれば、その時点の映像情報を破棄して、映像符号化を行わず、CPU時間を音声符号化部へ譲るものとするので、他のアプリケーションや、映像符号化部そのものが消費してしまったための計算機資源の不足による影響を、問題として認知されにくい映像のコマ落ちにとどめ、映像符号化のために音声途切れという事態に陥るのを回避することができる。

【0432】また、本実施の形態18の映像音声符号化装置によれば、映像符号化部1807の機能については、符号化負荷評価部1808が映像情報を出力すると、これを符号化すればよいものである。入力された映像を符号化するという単機能さえ有すれば、本実施の形態18による符号化装置の映像符号化部1807としてに適応可能である。すなわち、イメージ圧縮サブルーチン等の既存の映像符号化部を、内部の変更なしに、適用できるものである。このことは、圧縮サブルーチンがモジュール化されていて、後からアドオンできるような構成を持つ、汎用コンピュータ上の映像情報操作環境において、当該サブルーチンをそのまま応用して、映像音声の符号化に用いることができるので、ソフトウェア開発が効率的に行い得るという格別の効果を持つ。

【0433】なお、本実施の形態18において、マルチタスクを行う各プロセスのスイッチは、プロセス自身が計算機資源(CPU時間)を解放することにより行うものとしたが、本発明はこのような形態に限定されるもの

ではない。例えば、マルチタスクオペレーティングシステムが各プロセスに一定のCPU時間を与え、各プロセスがそのCPU時間を使い切ったら無条件に他のプロセスへスイッチする形態をとってもいい。この場合、映像符号化プロセスが、CPU時間を使い切る前に、音声符号化プロセスの進捗を監視し、必要があれば、自発的にCPU時間を解放するようにすれば、より効果的に計算機資源(CPU時間)の割り当てを行い、良好な符号化結果が得られるように図ることができる。

【0434】また、本実施の形態18では、符号化すべき映像情報(静止画像情報)を廃棄してしまうことで、映像符号化プロセスのCPU時間を解放している。すなわち、静止画像情報が、「0秒地点の静止画、1秒地点の静止画、2秒地点の静止画」と入力されたとき、必要があれば、1秒地点の静止画を落とし、「0秒地点の静止画、2秒地点の静止画」として符号化する。しかし、必ずしも最終的に出力される符号化情報が、コマ数の少ないものとなる必要はない。すなわち、静止画像情報の中に、どの地点(時間)での静止画像情報であるかを示すタイムスタンプを入れておき、映像符号化部が、そのタイムスタンプを確認することでコマ落としがあったかどうかを認識し、コマ落としがあった場合は、そのコマに相当する画像(前回の静止画と同一の画像や、前回と同一の画像であることを示す符号等)を出力すれば、最終的に出力される符号化映像情報は、額面上、コマ落としのない完全なものとなる。この手法をとれば、MPEG(Motion Picture Experts Group)規格など、映像情報のコマ数を既定値(1秒に30枚など)だけ保証しなければならない映像情報を出力するときでも、容易に対処できる。

【0435】実施の形態19、本発明の実施の形態19による映像音声符号化装置は、実施の形態18と同様に、汎用計算機等におけるソフトウェア処理において負担増大があった場合にも、音声の途切れを防ぐものであり、音声データの蓄積量を指標として、符号化に用いる予測処理の制御を行うものである。

【0436】図47は本発明の実施の形態19による映像音声符号化装置の概略構成を示す図である。図示するように、本実施の形態19による映像符号化装置は、ビデオカメラ1901、音声キャプチャ部1902、音声バッファリング部1903、音声符号化部1905、映像キャプチャ部1906、映像符号化部1923、および符号化負荷評価部1921から構成され、映像符号化部1923は、フレーム間予測処理部1924と、フレーム符号化部1925とを内包している。また、装置出力として符号化音声情報と、符号化映像情報とが出力されることは、実施の形態18と同様である。

【0437】同図において、符号化負荷評価部1921は、原音声バッファ量1904と、符号化負荷基準情報1910とに基づいて、符号化負荷評価情報1922を

計算により取得する。映像符号化部1923に内包されるフレーム間予測処理部1924は、映像の時間的冗長性を削減して圧縮符号化するために、静止画像情報間の動きベクトルを求め、動き補償を伴う予測符号化のために当該動きベクトルを出力する。映像符号化部1923に含まれるフレーム符号化部1925は、フレーム間予測処理部1924が出力した動きベクトルを用いて符号化を行い、符号化映像情報として出力する。

【0438】ビデオカメラ1901、音声キャプチャ部1902、音声バッファリング部1903、音声符号化部1905、および映像キャプチャ部1906については、実施の形態18の1801~1803、1805、および1806と同様であり、説明を省略する。

【0439】符号化負荷評価部1921による符号化負荷評価情報1922の計算にあたっては、原音声バッファ量104から計算した、音声バッファリング部193が持つバッファの空きの割合を示す評価情報と、符号化負荷基準情報110とを用いて両者を乗算して求めるものである。符号化負荷基準情報1910は、実施の形態18における場合と同様のものであり、本実施の形態19においても、固定的に「1」としておくものである。従って、本実施の形態19では、符号化負荷基準情報1910は考慮に入れることを要せず、評価情報として得られる、音声バッファリング部193が持つバッファの空きの割合を、そのまま符号化負荷評価情報1922とするものであり、バッファが空の場合は100%、一杯の場合は0%という値をとることとなる。

【0440】一般に、圧縮符号化にあたっては、1フレーム（1画面相当）の静止画像について、その空間的相関関係に基づいて圧縮を行うフレーム内符号化と、時間的に近接する、例えば連続するフレームの静止画像について、その時間的相関関係に基づいて圧縮を行うフレーム間符号化とがあり、フレーム内符号化が基本となるものではあるが、この両者を組み合わせると高圧縮率の符号化データが得られることとなる。フレーム間符号化を行うためには、フレームごとの動きを動きベクトルとして検出し、この動きベクトルを用いた動き補償を伴って予測画像を生成し、該予測画像と符号化対象である画像との差分データを圧縮するという手法が用いられる。

【0441】フレーム間予測処理部1924は、映像の時間的冗長性を削減して圧縮符号化するための予測画像生成処理に用いる、静止画像情報間の動きベクトルを求めるものである。本実施の形態19では、フレーム間予測処理部1924は、指定された割合だけ、予測処理を行うものである。すなわち、予測処理を行う最大の範囲を初期値として、処理を行う際に符号化負荷評価情報1922を入力し、初期値に対して、符号化負荷評価情報1922で示される割合だけ、フレーム間予測処理を行う。符号化負荷評価情報1922が100%であったら、初期値である最大の量だけフレーム間予測処理を行

い、得られた最適な動きベクトルを出力する。一方、符号化負荷評価情報1922が50%であったら、初期値の50%の量だけフレーム間予測処理を行い、その時点で求められた最適な動きベクトルを出力する。いずれの場合にも、フレーム符号化部1925は、出力された動きベクトルを用いて、符号化処理を行う。動きベクトルを求める処理は、処理量を増やせば増やすだけ最適な動きベクトルが求められることとなるので、このことにより予測画像と符号化対象画像との差が小さくなることから、効率的な圧縮が行えることとなる。一方、処理量を少なくすると最適なベクトルが得られず、圧縮率が下がる。なお、処理量を増大することなく圧縮率を維持することは、画質を犠牲にするならば可能である。

【0442】このように構成された、本実施の形態19による映像音声符号化装置の動作の概略は、以下のようになる。すなわち、当該映像音声符号化装置において、符号化負荷評価部1921は、音声バッファリング部1903が原音声情報を蓄積し、該蓄積量に応じて原音声バッファ量1904を更新した時点における原音声バッファ量1904と、符号化負荷基準情報1910とに従い、符号化負荷評価情報1922を出力する。また、映像符号化部1923は、映像キャプチャ部1906から出力された原映像情報を符号化して出力する。このとき、フレーム間予測処理部1924は、映像の時間的冗長性を削減して圧縮符号化するために、静止画像情報間の動きベクトルを求め、動き情報を用いて符号化を行う。これに伴い、フレーム符号化部1925は、フレーム間予測処理部1924が出力した動きベクトルを用いて符号化を行い、符号化映像情報として出力する。原音声情報の読み出しと符号化は、実施の形態18と同様に行われる。

【0443】以下に、ある映像音声に対しての、本実施の形態19による映像音声符号化装置による符号化処理の一例における動作を説明する。ここで、実施の形態18と同様に、映像音声符号化処理は、汎用計算機においてオペレーティングシステムの制御に従う映像符号化（符号化負荷評価部1921と映像符号化部1923の処理）と、音声符号化（音声符号化部1905の処理）との各タスクとしてなされるものであるとし、CPU時間の割り当てをされた各タスクが一連の処理を実行して計算機資源（CPU時間）を解放したとき、オペレーティングシステムは他のタスクにCPU時間の割り当てを行う、という制御をするものとする。

【0444】まず、実施の形態18と同様に、ビデオカメラ1901が映像音声情報を取り込み、アナログ音声情報とアナログ映像情報とに分けて出力する。そして、音声キャプチャ部1902は、ビデオカメラ1901から出力されたアナログ音声情報を入力し、デジタル原音声情報として出力する。音声バッファリング部1903は、原音声情報を蓄積し、蓄積量に応じて原音声バッ

ア量1904を更新する。一方、映像キャプチャ部1906は、ビデオカメラ1901から出力されたアナログ映像情報を入力し、デジタル原映像情報として出力する。

【0445】符号化負荷評価部1921は、その時点の原音声バッファ量104を確認する。ここでは、入力された原音声情報がバッファに30%ほど蓄積されていたので、符号化負荷評価情報1922は、70%とする。フレーム間予測処理部1924は符号化負荷評価情報1922を取得する。そして符号化負荷評価情報1922が70%であるので、初期値の70%だけフレーム間予測処理を行い、その中で最適な動きベクトルを得てこれをフレーム符号化部1925に出力する。フレーム符号化部1925は、動きベクトルを用い、映像情報を符号化し、符号化映像情報として出力し、映像符号化に割り当てられたCPU時間を解放する。

【0446】音声符号化部1905は、音声バッファリング部1903から一定量（原音声読み出し量）の原音声情報を、先（過去）に蓄積した分から読み出し、当該読み出した分の原音声情報を音声バッファリング部1903から削除して、原音声バッファ量1904を更新する。さらに、音声符号化部1905は、原音声情報を符号化する。本実施の形態18では、上記の原音声読み出し量について、最大バッファ量の30%とするので、上記の通り、音声バッファリング部1903に音声量が30%分蓄積されていたことから、そのすべてを読み出して符号化し、符号化が終了した時点でCPU時間を解放する。

【0447】CPU時間を割り当てられた符号化負荷評価部1921は、その時点の原音声バッファ量104を確認する。上記の音声符号化処理の直後であり、原音声バッファ量104に蓄積されている音声量が0%であったので、符号化負荷評価部1921は、符号化負荷評価情報1922を100%とする。そして、フレーム間予測処理部1924は、符号化負荷評価情報1922が100%であるので、初期値である最大量だけフレーム間予測処理を行い、最適な動きベクトルを取得し、フレーム符号化部1925は、該動きベクトルを用いて映像情報を符号化して、符号化映像情報を出力し、符号化が終了するとCPU時間を解放する。

【0448】図47において、ビデオカメラ1901より映像音声の取り込みが続いている間は、以上のように映像符号化、および音声符号化の処理が実行されることによって、当該取り込みにもなっている映像音声符号化が実行される。そして、映像音声の取り込みの終了後、符号化も終了する。

【0449】このように、本実施の形態19の映像音声符号化装置によれば、音声バッファリング部1903と、符号化負荷評価部1921と、フレーム間予測処理部1924、及びフレーム符号化部1925を内包する

映像符号化部1923とを備え、符号化負荷評価部1921が、映像符号化部1923が符号化を行う前に、その時点でのバッファに蓄積された未処理音声情報の量、すなわち原音声バッファ量1904を確認し、その量に応じて、フレーム間予測処理部1924における処理量を指示することにより、映像符号化で消費するCPU時間を制御するので、他のアプリケーションや、映像符号化部そのものが消費してしまったための計算機資源の不足による影響を、問題として認知されにくい一時的な圧縮率の低下や画質の低下にとどめ、映像符号化のために音声途切れという事態に陥るのを回避することができる。

【0450】また、本実施の形態19の映像音声符号化装置での、映像符号化部1923の機能については、符号化負荷評価部1921が符号化負荷評価情報1922を出力すると、これに対応して処理を行うものであるため、符号化負荷評価情報1922の入力と、該符号化負荷評価情報1922に対応した処理を行い得るものがある必要があり、実施の形態18のように、モジュール化したサブルーチンのそのままの応用がなし得るものではない。しかし、映像符号化の負荷を低減する際に、実施の形態18のように映像情報を破棄するのではなく、処理量を低減するものであるため、映像のコマ落ちは発生せず、実施の形態18と比較して、なめらかな動きの符号化映像情報が得られるという効果を持つ。

【0451】なお、本実施の形態19では、映像符号化で消費するCPU時間を制御するのに、フレーム間予測処理、すなわち最適な動きベクトルを求める計算の量を調節することを利用したが、本発明はこのような方式に限定されるものではない。例えば、色情報の符号化の処理の一部を省略するなど、他の映像符号化の処理を簡略化する方法をとることもできる。

【0452】また、本実施の形態19では、原音声バッファ量から求めたバッファの空き割合をそのまま符号化付加評価情報としているが、他の評価方法を用いることもできる。例えば、原音声バッファ量がある一定値を超えるまでは、符号化付加評価情報を100%とするが、一定値を超えてからは、50%、30%と減らしていく、という評価方法をとることもできる。

【0453】実施の形態20。本発明の実施の形態20による映像音声符号化装置は、実施の形態18と同様に、汎用計算機等におけるソフトウェア処理において負担増大があった場合にも、音声の途切れを防ぐものであり、音声データの蓄積量を指標として、符号化に用いる映像解像度の変更を行うものである。

【0454】図48は本発明の実施の形態20による映像音声符号化装置の概略構成を示す図である。図示するように、本実施の形態20による映像符号化装置は、ビデオカメラ2001、音声キャプチャ部2002、音声バッファリング部2003、音声符号化部2005、映

像キャプチャ部2031、映像符号化部2035、および符号化負荷評価部2032から構成され、映像符号化部2035は、映像符号化部本体2036と、解像度補正情報付加部2037とを内包している。また、装置出力として符号化音声情報と、符号化映像情報とが出力されることは、実施の形態18と同様である。

【0455】同図において、映像キャプチャ部2031は、実施の形態18、および19と同様に、アナログ映像情報より、静止画像の複数枚から構成されるデジタルの原映像情報を作成するものであるが、本実施の形態20では、後述する映像解像度情報2034を入力して、該入力した映像解像度情報解像度に対応する解像度を持つものとして、上記静止画像情報を作成するものである。映像キャプチャ部2031は実施の形態18と同様に、ビデオキャプチャボードで実現されるが、ここでは、当該ボードは解像度を指定し得るものであるとする。符号化負荷評価部2032は、符号化負荷評価情報2033を計算し、該計算した符号化負荷評価情報2033の値に従って、映像解像度情報2034を出力する。映像符号化部2035は、後述する映像符号化部本体2036と、解像度補正情報付加部2037とを内包し、映像キャプチャ部2031から出力された原映像情報を符号化し、符号化映像情報を出力する。映像符号化部本体2036は、映像符号化部2035に含まれ、実際の映像の符号化処理を行う。解像度補正情報付加部2037は、映像符号化部2035に含まれ、映像符号化部本体2036が出力した内部符号化映像情報に対して、解像度の情報を付加し当該映像音声符号化装置の装置出力となる符号化映像情報を作成する。

【0456】ビデオカメラ2001、音声キャプチャ部2002、音声バッファリング部2003、および音声符号化部2005については、実施の形態18の1801～1803、および1805と同様であり、説明を省略する。

【0457】上記の符号化負荷評価部2032による符号化負荷評価情報2033の計算にあたっては、原音声バッファ量2004の値に基づいて基本となる評価情報を計算し、それに符号化負荷基準情報2010の値を乗算して求める。本実施の形態20では、評価情報の計算にあたり、原音声バッファ量2004が音声バッファリング部2003の原音声蓄積可能量の半分を超えていれば、評価情報を0%とし、半分以下になれば、100%とするものである。また、符号化負荷基準情報2010は、実施の形態18における場合と同様であり、本実施の形態20においても、固定的に「1」としておくものである。従って、評価情報がそのまま符号化負荷評価情報2033となり、符号化負荷評価情報2033は、0%か100%という値をとる。符号化負荷評価部2032は、符号化負荷評価情報2033を用いて映像解像度情報2034を作成し、これを映像符号化部2035に

出力する。この際には、符号化負荷評価情報2033が100%である場合には、映像解像度情報2034を「幅320ピクセル、高さ240ピクセル」を示すものとして、また、0%であれば、「幅160ピクセル、高さ120ピクセル」を示すものとして出力する。

【0458】本実施の形態20においては、映像解像度情報2034の初期値として「幅320ピクセル、高さ240ピクセル」が設定されているものであるが、原音声バッファ量2004がバッファ量の最大の50%を超える場合に、上記のような符号化負荷評価部2032による演算において、「幅160ピクセル、高さ120ピクセル」に変化するものである。

【0459】このように構成された、本実施の形態20による映像音声符号化装置の動作の概略は、以下のようになる。すなわち、当該映像音声符号化装置において、映像キャプチャ部2031は、映像解像度情報2034を入力して、その解像度を持つ静止画像情報で構成されるデジタル原映像情報を、入力されたアナログ映像情報を変換することで作成し、出力する。また、符号化負荷評価部2032は、符号化負荷評価情報2033を計算し、この計算した符号化負荷評価情報2033の値に従って、映像解像度情報2034を出力する。そして、映像符号化部2035は、映像キャプチャ部2031から出力された原映像情報を符号化し出力する。このとき、映像符号化部本体2036は、実際の映像の符号化処理を行う。これに伴って、解像度補正情報付加部2037は、映像符号化部本体2035が出力した符号化映像情報に、解像度の情報を付加する。音声の扱いについては、実施の形態18と同様である。

【0460】以下に、ある映像音声に対しての、本実施の形態20による映像音声符号化装置による符号化処理の一例における動作を説明する。ここで、実施の形態18と同様に、映像音声符号化処理は、汎用計算機においてオペレーティングシステムの制御に従う映像符号化（符号化負荷評価部2032と映像符号化部2035の処理）と、音声符号化（音声符号化部2005の処理）との各タスクとしてなされるものであるとし、CPU時間の割り当てをされた各タスクが一連の処理を実行して計算機資源（CPU時間）を解放したとき、オペレーティングシステムは他のタスクにCPU時間の割り当てを行う、という制御をするものとする。

【0461】まず、実施の形態18と同様に、ビデオカメラ2001が映像音声情報を取り込み、アナログ音声情報とアナログ映像情報とに分けて出力する。そして、音声キャプチャ部2002は、ビデオカメラ2001から出力されたアナログ音声情報を入力し、デジタル原音声情報として出力する。音声バッファリング部2003は、原音声情報を蓄積し、蓄積量に応じて原音声バッファ量2004を更新する。

【0462】符号化負荷評価部2032は、その時点の

原音声バッファ量2004を確認する。上記入力された原音声情報がバッファに30%ほど蓄積されており、予め決められた基準値である50%を下回っているので、評価情報は100%となる。そして、上記のように値「1」である符号化負荷基準情報2010は、乗算処理しても結果に影響を与えないので、符号化負荷評価情報は100%となる。そこで、映像解像度情報2034は、「幅320ピクセル、高さ240ピクセル」となり、符号化負荷評価部2032は、この映像解像度情報2034を、映像キャプチャ部2031、および映像符号化部2035に出力する。

【0463】一方、映像キャプチャ部2031は、ビデオカメラ2001から出力されたアナログ映像情報を入力し、デジタル原映像情報として出力する。この際、映像キャプチャ部2031は、符号化負荷評価部2035から、映像解像度情報2034を入力される前であって、映像解像度情報2034の初期値である「幅320ピクセル、高さ240ピクセル」が用いられるものであり、映像キャプチャ部2031は、「幅320ピクセル、高さ240ピクセル」の静止画像情報からなるデジタル原映像情報を作成して出力する。

【0464】原映像情報は、映像符号化部2035に入力され、まず映像符号化部本体2036によって符号化処理をなされて、内部符号化映像情報が作成される。次いで、解像度補正情報付加部2037は、映像符号化部本体2036において作成された符号化映像情報に、「幅320ピクセル、高さ240ピクセル」を示す情報を付加し、当該映像音声符号化装置の装置出力となる符号化映像情報を作成して出力する。

【0465】ここで、実施の形態18での説明の場合と同様に、汎用計算機のオペレーティングシステム上で映像符号化、および音声符号化以外のタスクも実行されていたものとして、他のタスクにCPU時間の割り当てがなされ、「その他のプロセス」に制御が移って、CPU時間が消費される。実施の形態18に示したと同様に、映像音声のビデオカメラ2001による取り込みと、音声キャプチャ部2002、および映像キャプチャ部2031による処理とは、当該「その他のプロセス」とも概ね並行して行われるものであり、音声バッファリング部2003に音声量が90%分まで蓄積される。

【0466】その後、音声符号化処理が実行されると、音声符号化部2005は、音声バッファリング部2003から一定量の原音声情報を、先（過去）に蓄積した分から先に読み出し、その分の原音声情報を音声バッファリング部2003から削除し、原音声バッファ量2004を更新する。さらに、音声符号化部2005は、原音声情報を符号化する。本実施の形態20では、読み出して削除する一定量として30%であるものとし、音声符号化部2005は、上記のように90%まで蓄積された原音声情報のうちの30%を読み出して符号化し、

符号化を終了した時点でCPU時間を解放する。

【0467】映像符号化に再び処理が移ると、符号化負荷評価部2032は、その時点の原音声バッファ量2004を確認する。上記30%分の読み出しの直後で、まだ60%の原音声情報が蓄積されている。したがって、基準値の50%より大きい値となっているため、評価情報は0%となり、符号化負荷基準情報2010の「1」の乗算後にも値は変わらず、符号化負荷基準情報2033は0%となる。そこで、映像解像度情報2034は、「幅160ピクセル、高さ120ピクセル」となる。映像解像度情報2034は、先と同様に映像キャプチャ部2031と、映像符号化部2035とに出力される。

【0468】映像キャプチャ部2031は、アナログ映像情報を入力し、デジタル原映像情報として出力する。この際、映像解像度情報2034は、「幅160ピクセル、高さ120ピクセル」であるので、映像キャプチャ部2031は、「幅160ピクセル、高さ120ピクセル」の静止画像情報からなるデジタル原映像情報を出力する。

【0469】原映像情報は、映像符号化部2035に入力され、まず映像符号化部本体2036において符号化処理がされ、内部符号化映像情報として出力される。先の処理において、解像度が「幅320ピクセル、高さ240ピクセル」とことと比較すると、今回の処理では、映像の解像度が「幅160ピクセル、高さ120ピクセル」であるので、ピクセル数で表わされる情報量は4分の1となっている。従って、この符号化処理は、先の処理の4分の1の時間で終了する。解像度補正情報付加部2037は、映像符号化部本体2036から出力された内部符号化映像情報に対して、「幅160ピクセル、高さ120ピクセル」を示す情報を付加し、当該映像音声符号化装置の装置出力である符号化映像情報として出力する。

【0470】図48において、ビデオカメラ2001より映像音声の取り込みが続いている間は、以上のように映像符号化、および音声符号化の処理が実行されることによって、当該取り込みにもなった映像音声符号化が実行される。そして、映像音声の取り込みの終了後、符号化も終了する。

【0471】このように、本実施の形態20の映像音声符号化装置によれば、音声バッファリング部2003と、映像キャプチャ部2031と、符号化負荷評価部2032と、映像符号化部本体2036、および解像度補正情報付加部2037を内包する映像符号化部2035とを備えたことで、符号化負荷評価部2032が、音声バッファリング部2003に蓄積された未処理音声情報の量を確認し、その量に応じて、映像解像度情報2034を出力することによって入力される映像情報の解像度を制御し、これによって映像情報の情報量を制御する。よって、映像の符号化処理に消費されるCPU時間を制

御することができ、他のアプリケーションや、映像符号化部そのものが消費してしまった計算機資源の不足による影響を、問題として認知されにくい一時的な解像度の低下にとどめ、映像符号化のために音声途切れという事態に陥ることを回避することが可能となる。

【0472】なお、本実施の形態20では、映像解像度情報を変化させるのに、原音声バッファ量がある一定値を超えることを条件としているが、他の評価方法をとることもできる。例えば、原音声バッファ量にある係数を掛け、常にバッファ量に応じた解像度を設定することもでき、同様に音声の途切れ防止の効果が得られる。

【0473】実施の形態21. 本発明の実施の形態21による映像音声符号化装置は、実施の形態18と同様に、汎用計算機等におけるソフトウェア処理において負担増大があった場合にも、音声の途切れを防ぐものであり、音声データの処理量を指標として、映像情報の符号化処理停止を行うものである。

【0474】図49は本発明の実施の形態21による映像音声符号化装置の概略構成を示す図である。図示するように、本実施の形態21による映像符号化装置は、ビデオカメラ2101、音声キャプチャ部2102、音声バッファリング部2103、音声符号化部2142、映像キャプチャ部2106、映像符号化部2207、符号化負荷評価部2144、およびシステムタイマ2141から構成されている。また、装置出力として符号化音声情報と、符号化映像情報とが出力されることは、実施の形態18と同様である。

【0475】同図において、音声符号化部2142は、実施の形態1と同様に、音声バッファリング部2103に蓄積された原音声のうち、もっとも先（過去）に蓄積された原音声を取り出し、取り出した原音声を音声バッファリング部2103から削除し、原音声を圧縮符号化して符号化音声情報として出力する。これに加えて、本実施の形態21における音声符号化部2142は、これまでに取り出した原音声の総和である処理済み音声情報量2143を保持し、更新するものである。符号化負荷評価部2144は、後述する方式において、映像符号化の制御に用いる符号化負荷評価情報2145を計算により取得し、該取得した符号化負荷評価情報に対応して、原映像情報の符号化を実行するか否かを指示する。システムタイマ2141は、符号化の経過時間を測定する。

【0476】ビデオカメラ2101、音声キャプチャ部2102、音声バッファリング部2103、映像キャプチャ部2106、および映像符号化部2107については、実施の形態18の1801～1803、1806、および1807と同様であり、説明を省略する。

【0477】符号化負荷評価部1821による符号化負荷評価情報1922の計算にあたっては、まず、システムタイマ2141から求められる符号化の経過時間と、予め明らかである原音声情報の時間当たりの入力量とを

用いて原音声入力量を計算する。そして、計算により取得した原音声入力量と、音声符号化部2142が保持する処理済み音声情報量2143との差として、予測音声バッファ量を求める。次に、この求めた予測音声バッファ量を評価情報として用いて、実施の形態18と同様に符号化負荷基準情報2110との乗算処理によって符号化負荷評価情報2145を求める。本実施の形態21においても、符号化負荷基準情報2110は固定的に

「1」の値をとるものとし、予測音声バッファ量がすなわち、符号化負荷基準情報2110となるものである。そして、符号化負荷評価部2144は、符号化負荷評価情報2145の値を用いて、これが一定量を超えていなければ原映像情報を映像符号化部2107に出力して符号化を実行させ、一方一定量を超えている場合には、原映像情報を破棄することにより、符号化を実行させない。従って、本実施の形態21では、符号化負荷評価部は、予測バッファ量と一定量との比較を行うことになり、上記一定量としては、音声バッファリング部2103の最大バッファ量の50%とするものである。また、原音声情報の時間当たりの入力量は、10秒で音声バッファリング部2103のバッファが最大になるだけの量とする。

【0478】このように構成された、本実施の形態21による映像音声符号化装置の動作の概略は、以下のようになる。すなわち、当該映像音声符号化装置において、音声符号化部2142は、音声バッファリング部2103に蓄積された原音声のうち、もっとも先（過去）に蓄積された原音声を取り出し、取り出した原音声を音声バッファリング部2103から削除し、これまでに取り出した原音声の総和である処理済み音声情報量2143を更新し、原音声を圧縮符号化して符号化音声情報として出力する。そして、符号化負荷評価部2144は、システムタイマ2141から求められる符号化の経過時間と、予め明らかである原音声情報の時間当たりの入力量とで原音声入力量を計算し、この計算した原音声入力量と処理済み音声情報量2143との差である予測音声バッファ量を求め、この求めた予測音声バッファ量を用いて符号化負荷評価情報2145を求める。そして、この符号化負荷評価情報の値に応じて、映像符号化が制御される。音声の扱いについては、実施の形態18と同様である。

【0479】以下に、ある映像音声に対しての、本実施の形態21による映像音声符号化装置による符号化処理の一例における動作を説明する。ここで、実施の形態18と同様に、映像音声符号化処理は、汎用計算機においてオペレーティングシステムの制御に従う映像符号化（符号化負荷評価部2144と映像符号化部2107の処理）と、音声符号化（音声符号化部2142の処理）との各タスクとしてなされるものであるとし、CPU時間の割り当てをされた各タスクが一連の処理を実行して

計算機資源（CPU時間）を解放したとき、オペレーティングシステムは他のタスクにCPU時間の割り当てを行う、という制御をするものとする。

【0480】まず、実施の形態18と同様に、ビデオカメラ2101が映像音声情報を取り込み、アナログ音声情報とアナログ映像情報とに分けて出力する。そして、音声キャプチャ部2102は、ビデオカメラ2101から出力されたアナログ音声情報を入力し、デジタル原音声情報として出力する。音声バッファリング部2103は、原音声情報を蓄積し、蓄積量に応じて原音声バッファ量2104を更新する。一方、映像キャプチャ部2106は、ビデオカメラ2101から出力されたアナログ映像情報を入力し、デジタル原映像情報として出力する。

【0481】符号化負荷評価部2144は、一旦、映像キャプチャ部2144から出力された原映像情報を入力し、その時点において、予測音声バッファ量を確認する。このとき、システムタイマ2141を参照して得られる経過時刻は1秒であり、処理済み音声情報量2143はまだ「0」であるので、予測音声バッファ量は10%であって予め決められた基準値である50%を下回っており、符号化負荷基準情報2110は乗算処理にあたって考慮しないのでよい値「1」を有するので、符号化負荷評価情報2145は100%となる。そこで、符号化負荷評価部2144は、原映像情報を映像符号化部2107に入力し、映像符号化部2107はこの原映像情報に対して映像符号化処理を行い、符号化処理が終了した時点でCPU時間を解放する。

【0482】ここで、実施の形態18での説明の場合と同様に、汎用計算機のオペレーティングシステム上で映像符号化、および音声符号化以外のタスクも実行されていたものとして、他のタスクにCPU時間の割り当てがなされ、「その他のプロセス」に制御が移って、CPU時間が消費される。実施の形態18に示したと同様に、映像音声のビデオカメラ2101による取り込みと、音声キャプチャ部2102、および映像キャプチャ部2106による処理とは、当該「その他のプロセス」とも概ね並行して行われるものであり、音声バッファリング部2103に音声量が90%分まで蓄積される。

【0483】その後、音声符号化処理が実行されると、音声符号化部2142は、音声バッファリング部2103から一定量の原音声情報を、先（過去）に蓄積した分から先に読み出し、その分の原音声情報を音声バッファリング部2103から削除し、原音声バッファ量2104を更新する。さらに、音声符号化部2142は、原音声情報を符号化する。本実施の形態21では、読み出して削除する一定量として30%であるものとし、音声符号化部2105は、上記のように90%まで蓄積された原音声情報のうちの30%を読み出して符号化し、符号化を終了した時点で、30%分の量を自らの保持す

る処理済み音声情報量2143に加えて更新をし、CPU時間を解放する。

【0484】映像符号化に再び処理が移ると、符号化負荷評価部2144は、まずシステムタイマ2141を参照して、その時点の経過時間を確認する。「その他のプロセス」に移っていたため、経過時間は9秒である。次に音声符号化部2142の保持する処理済み音声情報量2143を参照すると、30%であった。このため、予測音声バッファ量は60%となり、基準値である50%を超えているので、評価情報は0%となり、符号化負荷基準情報2110の「1」を乗算して得られる符号化負荷評価情報も0%となる。そこで、符号化負荷評価部2144は、その時点の原映像情報を破棄し、すみやかにCPU時間を解放して、音声符号化が実行されるように図る。

【0485】図49において、ビデオカメラ2101より映像音声の取り込みが続いている間は、以上のように映像符号化、および音声符号化の処理が実行されることによって、当該取り込みにともなう映像音声符号化が実行される。そして、映像音声の取り込みの終了後、符号化も終了する。

【0486】このように、本実施の形態21による映像音声符号化装置によれば、システムタイマ2141と、処理済み音声情報量2143を保持する音声符号化部2142と、符号化音声負荷評価部2144とを備えたことで、符号化負荷評価部2144は、システムタイマ2141を参照して得られる経過時間と、音声符号化部2142を参照して得られる処理済み音声情報量2143とから予測バッファ量を計算し、この予測バッファ量を、原音声バッファ量2104の代替に用いて、映像符号化を制御するので、原音声バッファ量を用いて制御を行った実施の形態18と同様に、他のアプリケーションや、映像符号化部そのものが消費してしまった計算機資源の不足による、音声の途切れを防止することが可能となる。

【0487】また、本実施の形態21では、実施の形態18と異なり、原音声バッファ量2104がわからなくても、音声符号化部2142が処理した情報量を参照することで、現在蓄積されているであろう原音声バッファ量2104を予測することができるようになっているため、バッファ部がブラックボックスとなっている既存アプリケーションを使用する際にも、容易に対処できるものである。

【0488】なお、本実施の形態21では、状況に応じて映像符号化処理の停止を行う実施の形態18に準じた構成において、処理済み音声情報量に基づく予測バッファ量を指標とした映像符号化の制御を行うものとしたが、フレーム間予測符号化の処理量を制御する実施の形態19、および解像度を変更する実施の形態20に対して、予測バッファ量を指標とする本実施の形態21の手

法を応用することも可能であり、原音声バッファ量2104を知ることができない場合にも制御を行い得るとい同様の効果が得られる。

【0489】実施の形態22. 本発明の実施の形態20による映像音声符号化装置は、実施の形態18と同様に、汎用計算機等におけるソフトウェア処理において負担増大があった場合にも、音声の途切れを防ぐものであり、音声の符号化量を指標として、映像情報の符号化処理停止を行うものである。

【0490】図50は本発明の実施の形態22による映像音声符号化装置の概略構成を示す図である。図示するように、本実施の形態22による映像符号化装置は、ビデオカメラ2201、音声キャプチャ部2202、音声バッファリング部2203、音声符号化部2205、映像キャプチャ部2206、映像符号化部2207、符号化負荷評価部2253、およびシステムタイマ2251から構成されている。また、装置出力として符号化音声情報と、符号化映像情報とが出力されることは、実施の形態18と同様である。

【0491】同図において、符号化負荷評価部2253は、実施の形態21と同様に予測音声バッファ量を求め、これに基づいて符号化負荷評価情報を取得するものであるが、当該予測音声バッファ量を求める方法が実施の形態21による符号化装置とは異なるものである。本実施の形態22では、符号化負荷評価部2253は、音声符号化部2205から出力される符号化音声量2252を検出するものであり、この符号化音声量2252から得られる処理済み音声情報量2254を、実施の形態21における処理済み音声量2143の代わりに用いるものである。符号化音声情報は、上記の通りに、当該映像音声符号化装置の装置出力となり、伝送・記録等されるものであって、容易にその量を検出することができる。本実施の形態22の符号化負荷評価部2253においても、システムタイマ2251から経過時間を得る点、該経過時間と時間当たりの原音声入力量とから原音声入力量を取得する点、また、固定的に「1」の値とする符号化負荷基準情報2210を用いる点は実施の形態21の場合と同様である。

【0492】本実施の形態22による映像音声符号化装置は、上記のように符号化負荷評価部2253の機能が異なる点、および音声符号化部2205が処理済み音声量を保持しない点を除いては、実施の形態21による映像音声符号化装置と同様の構成となっているものである。従って、ビデオカメラ2201、音声キャプチャ部2202、音声バッファリング部2203、映像キャプチャ部2206、および映像符号化部2207については、実施の形態18の1801～1803、1806、および1807と同様であり、システムタイマ2251については実施の形態21と同様であるので説明を省略する。

【0493】本実施の形態22においても、実施の形態21と同様に、予測音声バッファ量と比較する一定量を最大バッファ量の50%、また、原音声の時間当たりの入力量を10秒で音声バッファリング部2203のバッファが最大になるだけの量とする。さらに、音声符号化部2205の圧縮率を10分の1とする。

【0494】このように構成された、本実施の形態21による映像音声符号化装置の動作の概略は、以下のようになる。すなわち、当該映像音声符号化装置の符号化負荷評価部2253は、システムタイマ2251から求められる符号化の経過時間と、予め明らかである原音声情報の時間当たりの入力量とを用いて原音声入力量を計算し、さらに、音声符号化部2205が出力した符号化音声情報の総量である符号化音声量2252から処理済み音声情報量2254を求め、さらに、原音声入力量と処理済み音声情報量2254との差である予測音声バッファ量を求め、この求めた予測音声バッファ量を用いて符号化負荷評価情報2209を求める。そして、符号化負荷評価部2253は、予測音声バッファ量が一定量より少なければ、原映像情報を映像符号化部2207に入力し、一定量以上であれば、原映像情報を破棄して映像符号化部2207の処理を中断し、計算機資源(CPU時間)を音声符号化部2205に明け渡す。

【0495】以下に、ある映像音声に対しての、本実施の形態22による映像音声符号化装置による符号化処理の一例における動作を説明する。ここで、実施の形態18と同様に、映像音声符号化処理は、汎用計算機においてオペレーティングシステムの制御に従う映像符号化(符号化負荷評価部2253と映像符号化部2207の処理)と、音声符号化(音声符号化部2205の処理)との各タスクとしてなされるものであるとし、CPU時間の割り当てをされた各タスクが一連の処理を実行して計算機資源(CPU時間)を解放したとき、オペレーティングシステムは他のタスクにCPU時間の割り当てを行う、という制御をするものとする。

【0496】まず、実施の形態18と同様に、ビデオカメラ2201が映像音声情報を取り込み、アナログ音声情報とアナログ映像情報とに分けて出力する。そして、音声キャプチャ部2202は、ビデオカメラ2201から出力されたアナログ音声情報を入力し、デジタル原音声情報として出力する。音声バッファリング部2203は、原音声情報を蓄積し、蓄積量に応じて原音声バッファ量2204を更新する。一方、映像キャプチャ部2206は、ビデオカメラ2201から出力されたアナログ映像情報を入力し、デジタル原映像情報として出力する。

【0497】符号化負荷評価部2244は、一旦、映像キャプチャ部2144から出力された原映像情報を入力し、その時点において、予測音声バッファ量を確認する。このとき、システムタイマ2251を参照して得ら

れる経過時刻は1秒であり、符号化音声量2252はまだ「0」であるので、予測音声バッファ量は10%であって予め決められた基準値である50%を下回っており、符号化負荷基準情報2110は乗算処理にあたって考慮しないでよい値「1」を有するので、符号化負荷評価情報2245は100%となる。そこで、符号化負荷評価部2253は、原映像情報を映像符号化部2207に入力し、映像符号化部2207はこの原映像情報に対して映像符号化処理を行い、符号化処理が終了した時点でCPU時間を解放する。

【0498】ここで、実施の形態18での説明の場合と同様に、汎用計算機のオペレーティングシステム上で映像符号化、および音声符号化以外のタスクも実行されていたものとして、他のタスクにCPU時間の割り当てがなされ、「その他のプロセス」に制御が移って、CPU時間が消費される。実施の形態18に示したと同様に、映像音声のビデオカメラ2201による取り込みと、音声キャプチャ部2202、および映像キャプチャ部2206による処理とは、当該「その他のプロセス」とも概ね並行して行われるものであり、音声バッファリング部2203に音声量が90%分まで蓄積される。

【0499】その後、音声符号化処理が実行されると、音声符号化部2205は、音声バッファリング部2203から一定量の原音声情報を、先（過去）に蓄積した分から先に読み出し、その分の原音声情報を音声バッファリング部2203から削除し、原音声バッファ量2204を更新する。さらに、音声符号化部2205は、原音声情報を符号化する。本実施の形態22では、読み出して削除する一定量として30%であるものとし、音声符号化部2205は、上記のように90%まで蓄積された原音声情報のうちの30%を読み出して符号化し、符号化を終了した時点で、CPU時間を解放する。

【0500】映像符号化に再び処理が移ると、符号化負荷評価部2244は、まずシステムタイマ2241を参照して、その時点の経過時間を確認する。「その他のプロセス」に移っていたため、経過時間は9秒である。次に音声符号化部2205から出力された符号化音声量が、音声バッファリング部103のバッファ量の3%であり、圧縮率が10分の1であることから、処理済み音声情報量2254は、30%と求められる。このため、予測音声バッファ量は60%となり、基準値である50%を超えているので、評価情報は0%となり、符号化負荷基準情報2110の「1」を乗算して得られる符号化負荷評価情報も0%となる。そこで、符号化負荷評価部2144は、その時点の原映像情報を破棄し、すみやかにCPU時間を解放して、音声符号化が実行されるように図る。

【0501】図50において、ビデオカメラ2201より映像音声の取り込みが続いている間は、以上のように映像符号化、および音声符号化の処理が実行されること

によって、当該取り込みにもなっている映像音声符号化が実行される。そして、映像音声の取り込みの終了後、符号化も終了する。

【0502】このように、本実施の形態22による映像音声符号化装置によれば、システムタイマ2251と、符号化音声量2252から処理済み音声情報量2254を取得する符号化音声負荷評価部2253とを備えたことで、符号化負荷評価部2253は、システムタイマ2251を参照して得られる経過時間と、符号化音声量2252を参照して得られる処理済み音声情報量2254とから予測バッファ量を計算し、この予測バッファ量を、原音声バッファ量2204の代替に用いて、映像符号化を制御するので、原音声バッファ量を用いて制御を行った実施の形態18と同様に、他のアプリケーションや、映像符号化部そのものが消費してしまった計算機資源の不足による、音声の途切れを防止することが可能となる。

【0503】また、本実施の形態22では、実施の形態18、および実施の形態21と異なり、原音声バッファ量2204と、音声符号化部2253における処理量とがわからなくても、音声符号化部2253が処理して出力した情報量を参照することで、現在蓄積されているであろう原音声バッファ量2204を予測することができるようになっているため、バッファ部に加えて、音声符号化部もブラックボックスとなっている既存アプリケーションを使用する際にも、容易に対処することができる。

【0504】なお、本実施の形態22では、状況に応じて映像符号化処理の停止を行う実施の形態18において、符号化音声量に基づく予測バッファ量を指標とした制御を行うものとしたが、フレーム間予測符号化の処理量を制御する実施の形態19、および解像度を変更する実施の形態20に対して、予測バッファ量を指標とする本手法を応用することも可能であり、同様の効果が得られる。

【0505】さて、ここまで説明してきた実施の形態18ないし実施の形態22は、すべて、平均的には符号化が可能なるものであることを前提として、瞬間的、または短期的な負荷増大による影響から、一時的に計算機資源が少なくなったときに、音声の途切れを防ぐためのものである。これらの実施の形態による映像音声符号化装置を、基本的に計算機能力の乏しいコンピュータシステム上で動作するソフトウェアにより実現する場合に応用することも可能ではあるが、かかる場合には、種々の条件と対象においての符号化のすべてにおいて好適であるとは言えない。

【0506】図51は実施の形態18ないし実施の形態22までの映像音声符号化装置を、基本的に計算機能力の乏しいコンピュータシステム上で実現したときの、音声バッファ量の推移を示す図である。同図に示すよう

10

20

30

40

50

に、全体的に映像符号化の負荷が大きすぎるため、映像符号化をしている間に、未処理の音声情報が増大してしまう。その結果、映像符号化が終了した段階で、音声符号化を優先する制御がなされるようになり、急激に音声符号化の処理が優先して実行され、その間映像の符号化処理は停止する。未処理の音声情報が少なくなった時点で、再び映像の符号化に戻ると、また未処理の音声情報が増大する。このようなことが繰り返されると、符号化された映像情報は、高質のものから、急激に低質のものとなり、また高質のものに戻る、といったことを繰り返し、結果として再生して利用する場合に鑑賞しがたいものとなってしまふ。

【0507】実施の形態18～22による映像符号化を行う場合の、かかる課題を解決するため、以下に説明する本発明の実施の形態23、および実施の形態24による映像符号化装置は、基本的に能力の乏しい計算機上で実現され、やはり音声の途切れを防止しながらも、上述のような映像の画質の大きな変動を抑制し得るものである。

【0508】実施の形態23. 本発明の実施の形態23による映像音声符号化装置は、高性能でない汎用計算機等において映像音声の符号化処理をソフトウェア処理によって行う場合に対応することを、符号化負荷基準情報の設定によってするものである。

【0509】図52は本発明の実施の形態23による映像音声符号化装置の概略構成を示す図である。図示するように、本実施の形態23による映像符号化装置は、ビデオカメラ2301、音声キャプチャ部2302、音声バッファリング部2303、音声符号化部2305、映像キャプチャ部2306、映像符号化部2307、符号化負荷評価部2308、システムタイマ2361、および符号化負荷基準設定部2362から構成されている。装置出力として符号化音声情報と、符号化映像情報とが出力されることは、実施の形態18と同様である。

【0510】同図において、システムタイマ2361は経過時間を計測する。符号化負荷基準設定部2362は、単位時間ごとの原音声バッファ量2304の変動を調査し、変動の度合いによって符号化負荷基準情報2363を設定する。符号化負荷評価部2308は、固定された値を有する符号化負荷基準情報ではなく、符号化負荷基準設定部2382によって設定された符号化負荷基準情報を用いて、符号化負荷評価情報2309を計算により取得する。本実施の形態23による映像音声符号化装置は、実施の形態18による映像符号化装置に、システムタイマ2361と、符号化負荷基準設定部2362とを追加した構成であり、ビデオカメラ2301、音声キャプチャ部2302、音声バッファリング部2303、音声符号化部2305、映像キャプチャ部2306、および映像符号化部2307については実施の形態18の1801～1807と同様であり、説明を省略す

る。

【0511】符号化負荷基準設定部2362は、原音声バッファ量2304が、一定値を超えた場合、または一定値より少なくなった場合に「1カウント」とするカウント動作を実行し、単位時間当たりに3カウントを超えたところで、符号化負荷基準情報2363を設定しなおすものとする。

【0512】符号化負荷基準情報2363は、音声バッファが空である場合に映像の処理をどの程度行うかを示す情報であり、例えば100%を表わす「1」の値であれば、音声バッファが空である場合に映像の処理を100%行い、50%を表わす「0.5」の値であれば、音声バッファが空である場合に映像の処理を50%行う、ということを示すものである。

【0513】本実施の形態23において、符号化負荷基準情報2363は、その初期値が「1」であるものとしているが、原音声バッファ量2304の変動により、符号化負荷基準設定部2362によるカウント動作において3カウントを超えると、再設定がなされることによりその値が「0.5」になる。

【0514】符号化負荷評価部2308は、この値を用いて符号化負荷情報2309を求めるので、符号化負荷情報2309の値は、符号化負荷基準情報2363が「1」のときは、実施の形態18と同じように、0%か100%の値をとるが、符号化負荷基準情報2363が「0.5」のときは、0%か50%の値となる。50%の値となったとき、符号化評価部2308は、その時点で入力された原映像情報を100%映像符号化部2307に入力するのではなく、50%だけ映像符号化部2307に入力する。従って、この場合、いわゆるフルフレーム(30fps)の処理をするのではなく、15fpsの処理を行うことになる。

【0515】このように構成された、本実施の形態23による映像音声符号化装置の動作の概略は、以下のようになる。すなわち、当該映像音声符号化装置において、符号化負荷基準設定部2362は、システムタイマ2361の計時出力に基づき、単位時間ごとの原音声バッファ量2304の変動を調査し、変動の度合いによって符号化負荷基準情報2363を設定する。符号化負荷評価部2308は、設定された符号化負荷基準情報2363を用いて、符号化負荷評価情報2309を計算し、この計算された符号化負荷評価情報2309の値に従って、原映像情報を映像符号化部2307に入力するか、原映像情報を破棄して映像符号化部2307の処理を中断し、計算機資源(CPU時間)を音声符号化部2305に明け渡すか、否かの動作を決定する。従って、映像符号化部2307において原映像情報の符号化処理が行われる場合にも、すべての原映像情報が処理されるとは限らず、状況に応じた割合での処理が行われることとなる。

【0516】以下に、ある映像音声に対しての、本実施の形態23による映像音声符号化装置による符号化処理の一例における動作を説明する。ここで、実施の形態18と同様に、映像音声符号化処理は、汎用計算機においてオペレーティングシステムの制御に従う映像符号化（符号化負荷評価部2308と映像符号化部2307の処理）と、音声符号化（音声符号化部2305の処理）との各タスクとしてなされるものであるとし、CPU時間の割り当てをされた各タスクが一連の処理を実行して計算機資源（CPU時間）を解放したとき、オペレーティングシステムは他のタスクにCPU時間の割り当てを行行、という制御をするものとする。また、当該映像音声符号化装置を実現する汎用計算機の基本的能力は、実施の形態18による符号化装置を実現する場合よりも低いものであるとする。

【0517】まず、実施の形態18と同様に、ビデオカメラ2301が映像音声情報を取り込み、アナログ音声情報とアナログ映像情報とに分けて出力する。そして、音声キャプチャ部2302は、ビデオカメラ2301から出力されたアナログ音声情報を入力し、デジタル原音声情報として出力する。音声バッファリング部2303は、原音声情報を蓄積し、蓄積量に応じて原音声バッファ量2304を更新する。一方、映像キャプチャ部2306は、ビデオカメラ2301から出力されたアナログ映像情報を入力し、デジタル原映像情報として出力する。

【0518】符号化負荷評価部2308は、一旦映像キャプチャ部2306から出力された原映像情報を入力し、その時点の原音声バッファ量2304と符号化負荷基準情報2363とを確認する。この時点で、原音声バッファ量2304は10%であり、予め決められた基準値である50%を下回っているので、評価値は100%となり、しかも符号化負荷基準情報2363は初期値の「1」であるので、符号化負荷基準情報2363は100%となる。そこで、符号化負荷評価部2308は、原映像情報のすべてのフレームを映像符号化部2307に入力し、映像符号化部2307は映像符号化処理を行い、映像符号化が終了した時点でCPU時間を解放する。

【0519】実施の形態18の場合と同様に、CPU制御に従う各タスクのプロセスと概ね並行して、ビデオカメラ2301と、音声キャプチャ部2302、および映像キャプチャ部2306を実現するキャプチャボードとの動作は行われ得るものであり、映像符号化処理にあたり、CPU時間の消費が大きかったことから、音声バッファリング部2303に90%まで原音声情報が蓄積される。

【0520】音声符号化部2305は、音声バッファリング部2303から一定量（原音声読み出し量）の原音声情報を、先（過去）に蓄積した分から読み出し、当該

読み出した分の原音声情報を音声バッファリング部2303から削除して、原音声バッファ量2304を更新する。さらに、音声符号化部2305は、原音声情報を符号化する。本実施の形態23では、上記の原音声読み出し量について、最大バッファ量の30%とするので、上記の通り、音声バッファリング部2303に原音声情報が90%分蓄積されていたことから、そのうち30%を読み出して、原音声バッファ量2304を60%に更新し、原音声情報の符号化処理を行い、符号化が終了した時点でCPU時間を解放する。また、原音声バッファ量2304を監視する符号化負荷基準設定部2362は、原音声バッファ量2304が60%であるので、既定値である50%を超えたことを認識する。

【0521】映像符号化に再び処理が移ると、符号化負荷評価部2308は、その時点の原音声バッファ量2304と符号化負荷基準情報2363とを確認する。この時点で、原音声バッファ量2304は60%であり、予め決められた基準値の50%以上となっていたので、評価値は0%となり、それに符号化負荷基準情報2363の「1」を乗算しても符号化負荷評価情報2309は0%となる。そこで、符号化負荷評価部2308は、この時点の原映像情報のすべてのフレームを廃棄し、すみやかにCPU時間を解放する。従って、この映像符号化の過程においては、原音声バッファ量2304は変わらない。

【0522】音声符号化部2305は、音声バッファリング部2303から一定量（原音声読み出し量）の原音声情報を、先（過去）に蓄積した分から読み出し、当該読み出した分の原音声情報を音声バッファリング部2303から削除して、原音声バッファ量2304を更新する。さらに、音声符号化部2305は、原音声情報を符号化する。原音声読み出し量は30%であり、音声バッファリング部2303に原音声情報が60%分蓄積されていたことから、そのうち30%を読み出して、原音声バッファ量2304を30%に更新し、原音声情報の符号化処理を行い、符号化が終了した時点でCPU時間を解放する。

【0523】原音声バッファ量2304を監視する符号化負荷基準設定部2362は、この時点で、原音声バッファ量2304が30%であるので、既定値である50%を下回り、前回既定値を超えたことと合わせて、原音声バッファ量2304の変動が1カウントであることを認識する。すなわち1カウント分のカウント動作が行われる。

【0524】このような過程が繰り返されることにより、原音声バッファ量2304の変動による、符号化音声基準設定部2362のカウント動作が3カウントに達すると、符号化負荷基準設定部2362は、符号化負荷基準情報2363を、初期値「1」から15fpsを指示する「0.5」に設定し直す。

【0525】映像符号化に再び処理が移ると、符号化負荷評価部2308は、その時点の原音声バッファ量2304と符号化負荷基準情報2363とを確認する。ここで、原音声バッファ量2304は30%であるので、評価値は100%となる。しかし、符号化負荷基準情報2363は「0.5」であるので、乗算して符号化負荷基準情報2363は50%となる。そこで、原映像情報のフレームのうち、半分を間引いて廃棄し、残りのフレームを映像符号化部2307に入力し、映像符号化処理を行い、終了した時点でCPU時間を解放する。

【0526】音声符号化部2305は、音声バッファリング部2303から一定量（本実施の形態23では30%）の原音声情報を、先（過去）に蓄積した分から先に読み出し、その分の原音声情報を音声バッファリング部2303から削除し、原音声バッファ量2304を更新する。さらに、音声符号化部2305は、原音声情報を符号化する。この過程では、先の映像符号化の過程に要する時間が、符号化負荷基準情報の再設定の前に比較して約半分で済んでいるので、音声バッファリング部2303に音声60%分蓄積されている。そのうちの30%を読み出し、30%とし、符号化が終了した時点でCPU時間を解放する。

【0527】映像符号化に再び処理が移ると、符号化負荷評価部2308は、その時点の原音声バッファ量2304と符号化負荷基準情報2363とを確認する。この時点で、原音声バッファ量2304は30%であり、既定値を下回っているが、符号化負荷基準情報2363は15fpsであるので、原映像情報のフレームのうち、半分を間引いて廃棄し、残りのフレームを映像符号化部2307に入力し、映像符号化処理を行い、終了した時点でCPU時間を解放する。

【0528】図52において、ビデオカメラ2301より映像音声の取り込みが続いている間は、以上のように映像符号化、および音声符号化のプロセスが実行されることによって、当該取り込みにもなっている映像音声符号化が実行される。そして、映像音声の取り込みの終了後、符号化も終了する。

【0529】図53は、このような本実施の形態23による映像音声符号化装置の符号化に際しての動作を長期の時間にわたって示す図である。同図に示すように、A区間では、音声バッファ量について図54に見られたような大きな変動があり、実施の形態18～22を性能の低い計算機で実行する場合に示したように、映像の符号化と、音声を優先する符号化とが繰り返されこととなっている。しかし、上記のように、この間に符号化負荷基準情報2363の再設定がなされるものであり、B区間では映像の符号化の負荷の基準が半分となったため、符号化の進行の度合いについてはバランスのよいものとなっている。

【0530】このように、本実施の形態23の映像音声

符号化装置によれば、実施の形態18による映像音声符号化装置にシステムタイマ2361と、符号化負荷基準設定部2362とを追加した構成としたことで、符号化負荷基準設定部2362は、原音声バッファ量2304の変動に対応して、符号化負荷基準情報2363の再設定を行い、符号化負荷評価部2306は符号化負荷基準情報2363を用いて符号化負荷評価情報2309を取得することにより、状況に応じて、原映像情報の符号化処理の割合を変更する。従って、負荷の変動により音声10がとぎれることを防止できるのに加えて、基本的に計算機能力の乏しいコンピュータシステム上で動作するソフトウェアにより映像音声を符号化する際において、そのシステムにおいて最適な映像負荷を自動的に設定し、それにより、高品質の映像と低品質の映像とが繰り返されてしまうといった事態に陥るのを回避することが可能となる。

【0531】なお、本実施の形態23では、状況に応じて映像符号化処理の停止を行う実施の形態18において、原音声情報蓄積量の変動に基づく、符号化負荷基準情報の変更による制御を行うものとしたが、フレーム間予測符号化の処理量を制御する実施の形態19、および解像度を変更する実施の形態20に対して、符号化負荷基準情報を変更する本手法を応用することも可能であり、本実施の形態23における場合と同様の効果が得られる。

【0532】例えば、実施の形態19に適用する際には、符号化負荷基準情報を、動きベクトルを計算する際、原音声バッファ量が「0」のときに、どのくらいの符号化を行うべきかを設定するものとし、原音声バッファ量の変動が大きい場合には、原音声バッファ量が「0」の場合でも、設定値の50%の量だけ動きベクトルを計算するとする等とする応用が可能である。従って、かかる応用によって、実施の形態19においても、高品質の映像と低品質の映像とが繰り返されてしまうといった事態に陥るのを回避することができる。

【0533】また、本実施の形態23は、初期値の符号化負荷基準情報で実際の符号化を行い、原音声バッファ量の変動の具合により、好適な符号化負荷基準情報を設定するというものであるが、設定して求めた符号化負荷基準情報は、ハードディスクなどの記憶装置に保存しておけば、次回符号化する際には、最初から、好適な符号化負荷基準情報で符号化が行える。すなわち、初回のみ図53に示すA区間のような好適な符号化の行えない期間が生じるが、次回からはB区間に示すような、バランスの良い符号化を実行し得るものとなる。

【0534】実施の形態24。本発明の実施の形態24による映像音声符号化装置は、高性能でない汎用計算機等において映像音声の符号化処理をソフトウェア処理によって行う場合に対応することを、符号化負荷基準情報の設定によってするものであり、当該設定の結果を利用

者に提示し得るものである。

【0535】図54は本発明の実施の形態24による映像音声符号化装置の概略構成を示す図である。図示するように、本実施の形態24による映像符号化装置は、ビデオカメラ2401、音声キャプチャ部2402、音声バッファリング部2403、音声符号化部2405、映像キャプチャ部2406、映像符号化部2407、符号化負荷評価部2408、システムタイマ2461、符号化負荷基準設定部2462、符号化負荷提示部2411、および負荷設定用標準映像音声出力部2412から構成されている。装置出力として符号化音声情報と、符号化映像情報とが出力されることは、実施の形態18と同様である。また、符号化負荷提示部2411は、モニタに対して出力を行うものである。

【0536】同図において、符号化負荷提示部2411は、符号化負荷基準情報2463の設定結果を、映像音声符号化装置の利用者に対して提示する。負荷設定用標準映像音声出力部2412は、コンピュータシステムの計算機能力に合わせて映像負荷を設定するため、標準的な映像情報、および音声情報を出力する。本実施の形態24による映像音声符号化装置は、実施の形態23による映像符号化装置に、符号化負荷提示部2411、および負荷設定用標準映像音声出力部2412を追加した構成であり、ビデオカメラ2401、音声キャプチャ部2402、音声バッファリング部2403、音声符号化部2405、映像キャプチャ部2406、および映像符号化部2407については実施の形態18の1801~1807と同様であり、また、システムタイマ2461と、符号化負荷基準設定部2462とは、実施の形態23と同様であるので、説明を省略する。

【0537】このように構成された、本実施の形態24による映像音声符号化装置の動作の概略は、以下のようになる。すなわち、当該映像音声符号化装置において、符号化負荷基準設定部2462は、システムタイマ2461の計時出力に基づき、単位時間ごとの原音声バッファ量2404の変動を調査し、変動の度合いによって符号化負荷基準情報2463を設定する。符号化負荷評価部2408は、設定された符号化負荷基準情報2463を用いて、符号化負荷評価情報2409を計算し、この計算された符号化負荷評価情報2409の値に従って、原映像情報を映像符号化部2407に入力するか、原映像情報を破棄して映像符号化部2407の処理を中断し、計算機資源(CPU時間)を音声符号化部2405に明け渡すか、否かの動作を決定する。従って、映像符号化部2407において原映像情報の符号化処理が行われる場合にも、すべての原映像情報が処理されるとは限らず、状況に応じた割合での処理が行われることとなる。

【0538】加えて、負荷設定用標準映像出力部2412は、標準映像音声情報2413に基づき、コンピュー

タシステムの計算機能力に合わせて映像負荷を設定するため、標準的な映像情報、および音声情報を出力する。また、符号化負荷提示部2411は、符号化負荷基準情報2463の設定結果を、映像音声符号化装置の利用者に対して提示する。

【0539】以下に、ある映像音声に対しての、本実施の形態24による映像音声符号化装置による符号化処理の一例における動作を説明する。ここで、実施の形態18と同様に、映像音声符号化処理は、汎用計算機においてオペレーティングシステムの制御に従う映像符号化(符号化負荷評価部2408と映像符号化部2407の処理)と、音声符号化(音声符号化部2405の処理)との各タスクとしてなされるものであるとし、CPU時間の割り当てをされた各タスクが一連の処理を実行して計算機資源(CPU時間)を解放したとき、オペレーティングシステムは他のタスクにCPU時間の割り当てを行う、という制御をするものとする。また、当該映像音声符号化装置を実現する汎用計算機の基本的能力は、実施の形態18による符号化装置を実現する場合よりも低いものであるとする。

【0540】まず、負荷設定用標準映像音声出力部2412が、実際の映像情報の符号化に先立ち、標準映像音声情報2413を出力する。実施の形態23においてとりこまれた映像音声情報と同様に、本実施の形態24による符号化装置は、標準映像音声情報2413を符号化し、符号化負荷基準設定部2462が、符号化負荷基準情報2463を設定する。符号化負荷提示部2411が、符号化負荷基準情報2463の内容を、モニタを通じて利用者に対して提示し、了解を得る。負荷設定用標準映像出力部2412が、出力する映像情報、および音声情報を、ビデオカメラ2401からの映像情報、および音声情報に切り替え、実施の形態23に示したような、通常の符号化を行う。

【0541】なお、本実施の形態24では、符号化負荷基準情報を設定するのに、予め用意した標準的な映像音声情報を用いるものとしたが、本発明はこのような方法に限定されるものではない。例えば、利用者が、符号化したい映像に合わせて、任意の映像音声情報を評価に用いることが可能である。また、本実施の形態24で用いる標準映像音声情報については、当該映像音声符号化装置における音声符号化が、音声情報の内容(有音・無音の別など)にかかわらずに、一定の符号化処理を行うものであれば、標準映像音声情報を構成する音声として無音のデータを用いることも可能である。

【0542】このように、本実施の形態24の映像音声符号化装置によれば、実施の形態23の映像音声符号化装置に、負荷設定用標準映像音声出力部2412と、符号化負荷表示部2411とを追加した構成としたことで、実際の符号化処理に先立ち、まず標準的な映像音声情報を符号化し、それにより、符号化負荷基準情報の設

定を行い、それを利用者に提示するようになっているため、利用者は、映像の品質の低下を納得した上で符号化処理を行える。すなわち、本実施の形態24で示したようなソフトウェアによる符号化装置は、様々なコンピュータシステムの上で動作可能である。したがって、計算機能力が元々高いものから、低いものまで、様々な環境の上で動作する。

【0543】このように、計算機能力の大小にかかわらず、映像の品質を犠牲にすることで音声の符号化を途切れなく行うというのが本発明の目的の1つであるが、本実施の形態24では、映像の品質をどれくらい犠牲にしたかを、利用者に示すことができる。これにより、利用者は、映像の品質の低下を、コンピュータシステムの計算機能力の不足によるものと認識でき、その対策として、動作周波数の向上であるとか、メインメモリの増設といったような、施策を講じることができる。したがって、本実施の形態24によれば、実施の形態23で示した効果に加え、利用者に当該コンピュータシステムについての状況を知らせることができるという格別の効果を得ることができる。

【0544】なお、本発明は上記各実施の形態に限定されるものではなく、例えば、上記各実施の形態においては、音声符号化部2405を、音声バッファリング部2403に蓄積された原音声情報を読み出し、この読み出した原音声情報を音声バッファリング部2403より削除した後、原音声情報を符号化処理し、符号化音声情報として出力するように構成したが、音声符号化部2405を、音声バッファリング部2403に蓄積された原音声情報を読み出し、この読み出した原音声情報を符号化処理し、符号化音声情報として出力した後、原音声情報を音声バッファリング部2403より削除するように構成してもよい。また、音声符号化部2405が音声バッファリング部2403に蓄積された原音声情報を読み出したことを検知した後、この読み出した原音声情報を音声バッファリング部2403より削除する削除部を別途設けてもよい。その他、本発明の請求の範囲内での種々の設計変更および修正を加え得ることが可能である。

【0545】また、実施の形態18~24に示した映像音声符号化方法については、該方法を実行し得る映像音声符号化プログラムを記録した記録媒体を用いて、パーソナルコンピュータやワークステーション等において、当該プログラムを実行することによって実現できるものである。

【0546】実施の形態25、本発明の実施の形態25による映像符号化方法は、実施の形態1と同様に、設定に応じて符号化パラメータを定めるものであり、入力画像データの有する解像度を与えられ、該解像度と、設定されたパラメータとに基づいて他のパラメータを決定するものである。

【0547】実施の形態1~4はいずれもフレームレ

ートを指定する場合であったが、本実施の形態25では、フレームレートを指定せず、なるべく高いフレームレートで符号化処理を実行し、再生画質の高い符号化データを得ようとするものである。また、本実施の形態25では、入力画像データの有する解像度与えられるものであり、当該与えられた解像度に対応して、符号化処理を行うものである。

【0548】図55は、本発明の実施の形態25による映像符号化装置の構成を示すブロック図である。図示するように、本実施の形態25による映像符号化装置は、符号化手段3001と、符号化パラメータ決定手段3002とから構成されており、符号化手段3001は、DCT処理手段3003、量子化手段3004、可変長符号化手段3005、ビットストリーム生成手段3006、逆量子化手段3007、逆DCT処理手段3008、および予測画像生成手段3009を、また、符号化パラメータ決定手段3002は動きベクトル(MV)検出範囲参照テーブル3010を内包している。

【0549】符号化手段3001は、映像がデジタル化された、一連の静止画像からなる映像データを入力画像データとして入力し、設定された符号化パラメータに従って符号化処理し、符号化データを出力する。入力画像データを構成する個々の静止画像データをフレーム画像と呼ぶ。また、符号化パラメータは、後述する符号化パラメータ決定手段3002から与えられるものであり、符号化タイプを示すパラメータと、動きベクトルの検出範囲を示すパラメータとが含まれている。符号化タイプを示すパラメータは、フレーム内符号化処理、または順方向予測符号化処理を示すものであり、符号化手段3001は、当該パラメータに従って、フレーム内符号化、または順方向予測符号化を行う。順方向予測符号化に用いる動きベクトルは、動きベクトルの検出範囲を示すパラメータによって指示される範囲内で、検出される。

【0550】符号化手段3001の内部のDCT手段3003、量子化手段3004、可変長符号化手段3005、ビットストリーム生成手段3006、逆量子化手段3007、および逆DCT手段3008については、実施の形態1における103~108と同様であり説明を省略する。

【0551】予測画像生成手段3009は、逆DCT処理手段3008が出力する逆DCT変換データを入力して、この逆DCT変換データと入力画像データとの間で動きベクトルの検出処理を行った後、予測画像を生成して予測画像データとして出力する。動きベクトルの検出は、上記のように動きベクトルの検出範囲を示すパラメータによって指示される範囲内で行われる。予測画像を用いたフレーム間符号化処理が行われる場合には、この予測画像データと入力画像データとの差分データがDCT手段3003に入力されることにより、符号化手段3001においては順方向予測符号化が行われることとな

る。

【0552】また、本実施の形態25による映像符号化装置では、符号化パラメータ決定手段3002は、入力画像データが有する解像度と、指定された符号化パターンとから、内包するMV検出範囲参照テーブル3010を用いてMV検出範囲を決定し、当該決定したMV検出範囲を示すパラメータを含む、上記符号化パラメータを符号化手段3001に出力する。

【0553】なお、本実施の形態25による映像符号化装置は、実施の形態1と同様に、パーソナルコンピュータ（PC）において処理制御装置（CPU）の制御により映像符号化プログラムが実行されることによって実現されるものとし、符号化処理の実行においては、実施の形態1において示した5つの条件に加えて以下の2つの条件が成立するものとする。

【0554】（6）順方向符号化処理が行われる場合、動きベクトルの検出範囲が「小さい」であるならば、処理時間はフレーム内符号化処理を行う場合の6倍となる。

【0555】（7）順方向符号化処理が行われる場合、動きベクトルの検出範囲が「大きい」であるならば、処理時間は検出範囲が「小さい」である場合の4倍となる。

【0556】ここで、本装置に搭載されるCPUの動作周波数は100MHzであり、符号化開始時に指定されるフレームレートは24フレーム/秒、符号化タイプの組み合わせとしての符号化パターンは、2フレームごとに「I」「P」を繰り返すパターン2「IP」を用いるものとする。ただし、フレーム内符号化を「I」、順方向予測符号化を「P」で表すものとする。

【0557】以上のような設定のもとに、上述のように構成された本実施の形態25による映像符号化装置の動作を以下に説明する。まず、符号化対象である映像はデジタル化され、一連のフレーム画像として当該符号化装置の符号化手段3001に入力される。図56は、符号化手段3001の動作を示すフローチャート図である。符号化手段3001の動作を、以下に、図56に従って説明する。なお、符号化パラメータ決定手段3002は、符号化開始時の最初のフレーム画像に対しては、符号化手段3001に対して必ずフレーム内符号化を指示するものとする。

【0558】ステップD01では、符号化パラメータ決定手段3002より入力された符号化パラメータについて判断がなされ、フレーム内符号化が指示されていた場合にはステップD02以降の処理が実行され、順方向予測符号化が指示されていた場合には、ステップD07以降の処理が実行される。

【0559】ステップD02以降が実行される場合は、ステップD02～ステップD05が実施の形態1におけるステップA02～ステップA05と同様に実行され

る。ステップD06では、符号化が終了しているか否かが判断され、符号化が終了したと判断されたならば処理は終了する。一方、符号化終了でなければ上記のステップD01に戻り、ステップD01の判断以降が実行される。

【0560】これに対して、ステップD01の判断により、ステップD07以降が実行される場合は次のようになる。まず、ステップD07で逆量子化手段3007は、量子化手段3004が直前のフレーム画像に対してすでに出力している量子化データを逆量子化し、逆量子化データを出力する。次いでステップD08では、逆DCT処理手段3008が、逆量子化データに対して、DCT処理手段3003が分割した8画素×8画素のブロックごとに、2次元離散コサイン変換の逆処理である2次元逆離散コサイン変換を実行し、逆DCT変換データを出力する。ステップD09において、予測画像生成手段3009は、逆DCT変換データに基づいて予測画像（未補償）を生成し、該生成した予測画像と入力画像データとに対して、符号化パラメータによって指示された範囲内で動きベクトル検出を行い、この動きベクトルを用いて、動き補償のされた予測画像を生成し出力する。

【0561】ステップD10でDCT処理手段3003は、入力されたフレーム画像と予測画像生成手段3009が出力した予測画像とを、それぞれ指示された解像度に基づき、8画素×8画素のブロックに分割し、分割したブロックごとに、入力されたフレーム画像のデータから予測画像のデータを差し引くことにより差分データを得る。そして、この差分データに対して、分割したブロックごとに2次元離散コサイン変換して、DCT変換データを出力する。DCT変換データが出力された後のステップD11～D14は、上記のステップD03からD06と同様に実行される。

【0562】このように、符号化手段3001では、入力されたフレーム画像ごとに、ステップD01の判定により、ステップD02～D06か、ステップD07～D14かの処理が行われることとなる。ステップD02～D06はフレーム内符号化であり、ステップD07～D14は直前のフレーム画像に対しての符号化結果を用いた予測画像に基づく順方向符号化処理が行われるものであり、この切り替えはステップD01の判定において、入力された符号化パラメータに従ってなされるものである。

【0563】（表15）は符号化パラメータ決定手段3002が内包するMV検出範囲参照テーブル3010を示す表である。また、図57は符号化パラメータ決定手段3002の動作を示すフローチャート図である。以下に、符号化パラメータを決定して、符号化手段3001に出力する符号化パラメータ決定手段3002の動作を、表1を参照し、図57のフローに従って説明する。

【0564】

【表15】

入力		出力
解像度	符号化パターン	MV検出範囲
160×120	I P	小
80×64	I P	大

【0565】(表15)に示すMV検出範囲参照テーブル3010は、符号化に先立ち予め作成しておかれるものである。テーブル作成は、後述する条件を考慮した上で、例えば経験的知識に基づいて、あるいは実験符号化やシミュレーション等の結果を用いて、することが10
 できる。(表1)の「入力」の欄は入力画像データの有する解像度と、指示されるパラメータとを、また「出力」の欄は入力に対応して決定されるパラメータを示している。同表に示すように、本実施の形態25では入力画像データの解像度と符号化パターンとに対応して、MV検出範囲が決定される。符号化パターンについては固定的に「I P」とされているものであり、パターン「I P」は、2フレームごとにフレーム内符号化(I)と順方向予測符号化(P)とを繰り返すことを意味する。入力画像の解像度としては、「160×120」と「80
 ×64」とのいずれかが指示されるものとする。

【0566】参照テーブルの作成は、次の条件を考慮して行われる。第一に、動きベクトル検出範囲が大きくなると、処理量が多くなること、第二に入力画像の有する解像度が高い場合は、低解像度である場合と比べて処理量が多くなることである。

【0567】これらの条件を考慮し、指定された入力画*

入力		出力	符号化結果
解像度	符号化パターン	MV検出範囲	フレームレート
160×120	I P	小	27.4
80×64	I P	大	27.4

【0573】(表16)は、指示される条件に対して、本実施の形態25の符号化装置において決定されるMV検出範囲(決定されるパラメータ)と、それらのパラメータを用いた符号化処理の結果として得られたフレームレート(符号化結果)とを示している。(表16)に示す符号化結果の数値については、符号化パターン「I P」において、解像度を160×120とした場合に27.4フレーム/秒で処理できることに基づいて、その他の場合のフレームレートが算出されている。符号化パ
 40
 ターンがI Pで解像度が80×64、動きベクトルの検*

入力		出力	符号化結果
解像度	符号化パターン	MV検出範囲	フレームレート
160×120	I P	大	6.9
160×120	I P	小	27.4
80×64	I P	大	27.4
80×64	I P	小	109.7

【0576】(表17)においても(表16)の場合と同様の算出がなされており、符号化パターン「I P」において、解像度が160×120、検出範囲が「小さ
 50
 い」

* 像の解像度に対して、できるだけ大きな範囲で検出を行い、高圧縮率の符号化データを得られるように、MV検出範囲参照テーブル3010は作成されるものである。

【0568】まず、図57のフローのステップE01において、符号化パラメータ決定手段3002は、指定された入力画像の解像度と、符号化パターン(I P)とから、MV検出範囲参照テーブル3010を参照して、予測符号化における動きベクトルの検出範囲を決定する。

【0569】次いでステップE02では、符号化パラメータ決定手段3002は、符号化手段3001に対して、ステップE01で決定したMV検出範囲を指示するとともに、指定された符号化パターンを実現できるように、処理対象であるフレーム画像に用いるべき符号化タイプ(IもしくはP)を指示する。

【0570】その後、ステップE03では符号化が終了したか否かが判定され、符号化が終了したと判定されたならば処理は終了する。一方、終了でなければ、ステップE02に戻ることによって、符号化手段3001に対する符号化パラメータ出力が繰り返される。

【0571】符号化手段3001と、符号化パラメータ決定手段3002との以上のような動作によって、符号化が実行されるが、(表16)は、本実施の形態25による映像符号化装置において符号化を行なった結果を示す表である。

【0572】

【表16】

* 出範囲が「大きい」の場合のフレームレートは、検出範囲が大きい場合の処理に小さい場合の処理の4倍の時間を要することと、解像度が約1/4の場合は約1/4の時間で処理できることから、約27.4フレーム/秒と算出できる。

【0574】比較のため、(表17)に従来の技術による映像符号化装置を用いて符号化を行なった場合の動作結果を示す。

【0575】

【表17】

い」の場合に27.4フレーム/秒で処理できることに基づいて、その他の場合のフレームレートが算出されている。

【0577】従来の技術による映像符号化装置では、符号化結果として得られるフレームレートを考慮せずに、動きベクトル検出範囲を決定していたものである。従って、符号化処理の結果として得られるフレームレートが十分高くなるように設定することが困難なことが多かった。これに比べ、本実施の形態25の映像符号化装置においては、符号化結果であるフレームレートを考慮して、指定された符号化タイプ（パターン）と、入力される画像が有する解像度とに応じて動きベクトルの検出範囲を決定することで、表16と表17との対比において示されるように、できるだけ高いフレームレートを実現しつつ、より高圧縮率の符号化データを得られるように動きベクトルの検出範囲を設定しての符号化が実行されていることがわかる。

【0578】このように、本実施の形態25による映像符号化装置によれば、符号化手段3001と、動きベクトル検出範囲参照テーブル3010を内包した符号化パラメータ決定手段3002とを備えたことで、符号化パラメータ決定手段3002は、指定された符号化タイプと入力される画像が有する解像度に対応して、動きベクトルの検出範囲を決定して、符号化パラメータを符号化手段3001に出力し、符号化手段3001はこの符号化パラメータに応じて符号化の処理を行うので、要求される条件を実現しつつ、より高圧縮率の符号化データの得られる符号化を行うことが可能となる。

【0579】なお、本実施の形態25による映像符号化装置では、指定された符号化パターンと入力画像の有する解像度に対応して動きベクトルの検出範囲を決定するものとしたが、同様の参照テーブルを用いることによって、入力画像の有する解像度に対応してフィルタリングの有無を決定することも可能であり、設定された条件下で、より高画質の符号化結果の得られる処理をすることが可能となる。

【0580】なお、実施の形態25に示した映像符号化方法については、実施の形態1と同様に、該方法を実行し得る映像音声符号化プログラムを記録した記録媒体を用いて、パーソナルコンピュータやワークステーション等において、当該プログラムを実行することによって実現できるものである。

【0581】また、実施の形態1～25のいずれについても、当該符号化プログラムを記録する媒体としては、フロッピーディスク、CD-ROM、光磁気ディスク、相変換型光ディスク等、当該符号化プログラムを記録することができ、パーソナルコンピュータ等の汎用計算機で読み出して実行できるものであれば使用可能である。又、ネットワークにより接続する、他のコンピュータが管理する記憶装置に記録した当該プログラムを、ネットワークを介して読み出し、当該読み出したコンピュータにおいて実行する運用形態をとることも可能である。

【0582】

【発明の効果】請求項1の映像符号化方法によれば、映像がデジタル化された、複数の静止画像情報からなる原映像情報に対して、上記静止画像情報の1つまたは複数、後述する符号化パラメータに従って符号化する映像符号化ステップと、原映像情報の有する解像度、符号化によって得られる符号化データを再生する際に要求されるフレームレート、上記映像符号化ステップを実行する符号化装置の処理能力を示す処理性能、または上記映像符号化ステップにおける符号化処理の処理量に影響する1つ、もしくは複数の符号化パラメータのうちいずれか1つ以上に基づいて、1つ以上の上記符号化パラメータを決定する符号化パラメータ決定ステップとを実行するので、与えられたフレームレートにおいて、高解像度の、又は高圧縮率の符号化結果の得られる符号化パラメータを設定することで、装置資源を活用して、良好な符号化結果を得ることが可能となる。

【0583】また、請求項2の映像符号化方法によれば、請求項1の方法において、当該映像符号化方法の、上記映像符号化ステップを実行する符号化装置の処理能力を判断して、判断結果を出力する処理能力判断ステップをさらに実行するので、処理能力に応じた良好な符号化結果を得ることが可能となる。

【0584】また、請求項3の映像符号化方法によれば、請求項1または2の方法において、上記符号化パラメータは、上記原映像情報に対して行う符号化処理における解像度、フレーム内符号化、もしくは予測符号化を示す符号化タイプ、または上記予測符号化に用いる動きベクトルを検出する際の検出範囲のうち1つ以上を含むものとするので、設定に応じてこれらのパラメータを決定して、装置資源を活用して、良好な符号化結果を得ることが可能となる。

【0585】また、請求項4の映像符号化方法によれば、請求項3の方法において、上記処理能力判断ステップでは、当該映像符号化方法の有する制御装置の種類に基づいて上記判断を行うので、当該符号化装置のハードウェア能力を判断し、その能力に応じて符号化パラメータを設定することで、装置資源を活用して良好な符号化結果を得ることが可能となる。

【0586】また、請求項5の映像符号化方法によれば、請求項3の方法において、上記処理能力判断ステップでは、上記符号化ステップにおける符号化処理の所要時間に基づいて上記判断を行うので、上記所要時間が示す、当該符号化装置の符号化処理能力を判断し、その能力に応じて符号化パラメータを設定することで、装置資源を活用して良好な符号化結果を得ることが可能となる。

【0587】また、請求項6の映像符号化方法によれば、請求項3の方法において、上記処理能力判断ステップでは、上記入力される原映像情報を一時蓄積し、該蓄積にあたっては、上記原映像情報を構成する一連の静止

画像情報を順次保存していくとともに、上記符号化ステップにおいて読み出されて、上記符号化処理が行われた静止画像情報を順次廃棄する映像バッファリングステップと、上記映像バッファリングステップにおける上記一連の静止画像情報の保存を、上記与えられたフレームレートに基づいて決定される一定のフレームレートにおいて行うように制御するフレームレート制御ステップとを実行し、上記映像バッファリングステップにおいて一時蓄積された上記原映像情報の蓄積量に基づいて上記判断を行うので、上記蓄積量が示す、当該符号化装置の符号化処理能力を判断し、その能力に応じて符号化パラメータを設定することで、装置資源を活用して良好な符号化結果を得ることが可能となる。

【0588】また、請求項7の音声符号化方法によれば、音声に対して、帯域分割符号化方式により符号化を行う音声符号化方法において、符号化処理に用いる数値である、設定周波数 f_s と、変換定数 n とを記憶する記憶ステップと、符号化の対象である音声を入力する音声入力ステップと、上記記憶した設定周波数 f_s に基づいて決定されるサンプリング周波数を用いて、サンプリング音声データを作成する入力音声サンプリングステップと、上記設定周波数 f_s をサンプリング周波数として用いた場合に得られるサンプリング音声データの個数を m 個とし、上記変換定数 n に基づいて定められる数を m' として、 m' 個のサンプリング音声データを含む、 m 個の音声データからなる変換音声データを出力する音声データ変換ステップと、上記変換音声データを、帯域分割して M 個の帯域信号を得る帯域分割ステップと、上記記憶した設定周波数 f_s と変換定数 n とから得られる周波数 $f_s/2n$ を制限周波数として、上記帯域信号のうち、制限周波数以下の帯域信号にのみ符号化ビットを割り当てる符号化ビット割り当てステップと、上記割り当てた符号化ビットに基づいて量子化を行う量子化ステップと、上記量子化したデータを符号化データとして出力する符号化ステップと、上記出力される符号化データを記録する符号化データ記録ステップとを実行するので、変換定数 n の設定によって、処理負担を軽減し、音声取り込みにともなったリアルタイム符号化処理を行って、当該符号化装置の性能に応じた音質の符号化結果を得ることが可能となる。

【0589】また、請求項8の音声符号化方法によれば、請求項7の方法において、上記入力音声サンプリングステップでは、上記記憶した設定周波数 f_s をサンプリング周波数として、上記入力された音声のサンプリング処理により、 m 個のサンプリング音声データを作成するものであり、上記音声データ変換ステップでは、上記 m 個のサンプリング音声データより、 $(n-1)$ 個おきにサンプリング音声データを抽出し、2つの隣接する上記抽出したサンプリング音声データの間に、 $(n-1)$ 個の音声データを挿入して、 m 個の変換音声データに変

換するので、処理負担を軽減し得る変換音声データを得ることによって、音声取り込みにともなったリアルタイム符号化処理を行って、当該符号化装置の性能に応じた音質の符号化結果を得ることが可能となる。

【0590】また、請求項9の音声符号化方法によれば、請求項8の方法において、上記音声データ変換ステップでは、上記抽出したサンプリング音声データがそれぞれ n 個ずつ連続する変換音声データを作成することで、上記の処理負担を軽減し得る変換音声データを、容易に得ることが可能となる。

【0591】また、請求項10の音声符号化方法によれば、請求項7の方法において、上記入力音声サンプリングステップでは、上記記憶した設定周波数 f_s と変換定数 n とから得られる周波数 f_s/n をサンプリング周波数として、上記入力された音声のサンプリング処理により、 m/n 個のサンプリング音声データを作成するものであり、上記音声データ変換ステップでは、上記サンプリング音声データに基づき、2つの隣接するサンプリング音声データの間に $(n-1)$ 個の音声データを挿入して、 m 個の変換音声データに変換するので、変換定数の設定によって音声入力時のサンプリング周波数を制御することにより、処理負担を軽減することで、音声取り込みにともなったリアルタイム符号化処理を行って、当該符号化装置の性能に応じた音質の符号化結果を得ることに加えて、サンプリングデータ量が低減することによる、データ一時蓄積のためのバッファ使用量の低減をも図ることが可能となる。

【0592】また、請求項11の音声符号化方法によれば、請求項10の方法において、上記音声データ変換ステップでは、上記 m/n 個のサンプリング音声データが、それぞれ n 個ずつ連続する変換音声データを作成することで、上記の処理負担を軽減し得る変換音声データを、容易に得ることが可能となる。

【0593】また、請求項12の音声符号化方法によれば、請求項7ないし11のいずれかの方法において、上記サンプリング音声データを、入力バッファに一時的に保持する音声バッファリングステップと、上記入力バッファのデータ量を調べて、これを予め設定した値と比較し、上記比較の結果に基づいて、上記レジスタに記憶された上記変換定数 n の値を変更する入力バッファ監視ステップとを実行し、上記入力音声サンプリングステップでは、上記サンプリング音声データを上記入力バッファに書き込むものであり、上記音声データ変換ステップでは、上記入力バッファよりサンプリング音声データを読み出して、これを上記変換するので、一時蓄積されるデータ量を指標として、その時点での当該符号化装置の処理能力を判断し、その結果に応じて変換定数 n の数値を変更することによって、当該装置の処理能力の変動に対応して、音声取り込みにともなったリアルタイム符号化処理を行うことが可能となる。

【0594】また、請求項13の音声符号化方法によれば、請求項7ないし11のいずれかの方法において、上記符号化ステップにおいて出力される単位時間当たりの符号化データ量を調べて、これを予め設定した値と比較し、上記比較の結果に基づいて、上記レジスタに記憶された上記変換定数 n の値を変更する符号化データ監視ステップを実行するので、符号化データ量を指標として、その時点での当該符号化装置の処理能力を判断し、その結果に応じて変換定数 n の数値を変更することによって、当該装置の処理能力の変動に対応して、音声取り込みにもなったリアルタイム符号化処理を行うことが可能となる。

【0595】また、請求項14の音声符号化方法によれば、音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化方法において、上記符号化に用いる制御定数を記憶する制御定数記憶ステップと、入力音声をサンプリング処理して、サンプリングデータを出力するサンプリングステップと、上記サンプリングステップで得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割ステップと、上記帯域分割ステップで得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当てステップと、上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化ステップと、上記量子化ステップで得られた量子化値に基づき、符号化データを出力する符号化ステップと、上記記憶した制御定数に基づいて、上記帯域分割ステップ、上記符号化ビット割り当てステップ、上記量子化ステップ、および上記符号化ステップにおけるデータ処理を制御する符号化処理制御ステップとを実行するので、制御定数の設定によって、処理負担を軽減し、音声取り込みにもなったリアルタイム符号化処理を行って、当該符号化装置の性能に応じた音質の符号化結果を得ることが可能となる。

【0596】また、請求項15の音声符号化方法によれば、請求項14の方法において、上記制御定数記憶ステップでは、上記制御定数として、単位期間判定定数 k を単位期間判定定数レジスタに記憶するものであり、上記符号化処理制御ステップは、上記帯域分割ステップでの1回の帯域分割処理で対象とするサンプリングデータ数を p とし、 p 個のサンプリングデータに相当する時間を単位期間として、上記出力されるサンプリングデータの p 個ごとに、相当する単位期間が符号化対象期間であるか符号化対象外期間であるかの判定を、上記記憶した単位期間判定定数に基づいて行い、上記単位期間が上記符号化対象期間と判定されたときのみ、該単位期間のサンプリングデータが上記帯域分割ステップに出力されるよう制御し、上記単位期間が上記符号化対象外期間と判定されたときは、上記符号化ステップにおいて、予め記憶した固定的符号化データを符号化データとして出力する

よう制御する判定制御ステップであるので、対象期間ごとに区切られたサンプリングデータに対して、符号化処理対象期間のもののみに対して符号化処理を実行し、他は符号化処理を実行せず、固定的符号化データを用いることで、処理負担を軽減し、音声取り込みにもなったリアルタイム符号化処理を行って、当該符号化装置の性能に応じた音質の符号化結果を得ることが可能となる。

【0597】また、請求項16の音声符号化方法によれば、請求項15の方法において、上記判定制御ステップでは、 i 番目の単位期間を t_i として、上記記憶した単位期間判定定数 k と任意の整数 n とから $i = n \times k + 1$ が成立するとき、上記単位期間 t_i が上記符号化対象期間であると判定するので、単位期間鑑定定数 k の設定に応じて上記符号化対象期間を定め、処理負担の軽減を図ることが可能となる。

【0598】また、請求項17の音声符号化方法によれば、請求項14の方法において、上記制御定数記憶ステップでは、上記制御定数として、演算処理判定定数 q を演算処理判定定数レジスタに記憶するものであり、上記符号化処理制御ステップは、上記帯域分割ステップに内包され、上記記憶した演算処理判定定数 q に基づいて、上記帯域分割ステップにおける演算処理を途中で打ち切るように制御する演算処理中止ステップであるので、演算処理判定定数 q の設定に応じて、帯域分割ステップにおける演算処理の一部を省略することによって、処理負担の軽減を図ることが可能となる。

【0599】また、請求項18の音声符号化方法によれば、請求項17の方法において、上記演算処理中止ステップでは、上記帯域分割ステップにおける基本低域通過フィルタの演算処理を、該フィルタの両端ステップ分については途中で打ち切るように制御することで、上記の演算処理の省略をし、処理負担の軽減を図ることが可能となる。

【0600】また、請求項19の音声符号化方法によれば、請求項14の方法において、上記制御定数記憶ステップでは、上記制御定数として、帯域選択定数 r を帯域選択定数レジスタに記憶するものであり、上記符号化処理制御ステップは、上記帯域分割ステップが出力する帯域信号データのうち、上記記憶した帯域選択定数 r に基づいて選択したもののみに対して、上記符号化ビット割り当てステップと上記量子化ステップとにおける処理を実行するよう制御する帯域間引きステップであるので、帯域選択定数 r の設定に応じて、帯域信号データの一部に対して、後段の処理を省略することによって、処理負担の軽減を図ることが可能となる。

【0601】また、請求項20の音声符号化方法によれば、請求項19の方法において、上記帯域間引きステップでは、上記帯域分割ステップで得られた M 個の帯域信号データ出力から、上記記憶した帯域選択定数である r 個おきに帯域信号データを選択することによって、上記

の帯域信号データの選択を実行し、処理負担の軽減を図ることが可能となる。

【0602】また、請求項21の音声符号化方法によれば、請求項14ないし20のいずれかの方法において、音声符号化におけるデータ処理の状況を取得し、該取得した状況に応じて、上記記憶した上記制御定数の値を変更する処理状況監視ステップを実行するので、データ処理の状況を指標として、その時点での当該符号化装置の処理能力を判断し、その結果に応じて制御定数の数値を変更することによって、当該装置の処理能力の変動に対応して、音声取り込みにともなったリアルタイム符号化処理を行うことが可能となる。

【0603】また、請求項22の音声符号化方法によれば、請求項21の方法において、上記処理状況監視ステップでは、サンプリングデータを入力バッファに一時蓄積する音声バッファリングステップと、上記入力バッファに保持されるデータの量を予め設定した値と比較し、上記比較の結果に基づいて上記制御定数変更を行う入力監視ステップとを実行することで、一時蓄積量を指標として、その時点での当該符号化装置の処理能力を判断し、その結果に応じて変換定数nの数値を変更することによって、当該装置の処理能力の変動に対応して、音声取り込みにともなったリアルタイム符号化処理を行うことが可能となる。

【0604】また、請求項23の音声符号化方法によれば、請求項21の方法において、上記処理状況監視ステップは、上記符号化ステップにおいて単位時間当たり出力される上記符号化データの量を、予め設定した値と比較し、上記比較の結果に基づいて上記制御定数の値を変更する符号化監視ステップであるものとしたことで、符号化データ量を指標として、その時点での当該符号化装置の処理能力を判断し、その結果に応じて変換定数nの数値を変更することによって、当該装置の処理能力の変動に対応して、音声取り込みにともなったリアルタイム符号化処理を行うことが可能となる。

【0605】また、請求項24の音声符号化方法によれば、音声デジタル化された原音声情報に対して、帯域分割符号化方式を用いて符号化を行う音声符号化方法において、入力音声をサンプリング処理して、サンプリングデータを出力するサンプリングステップと、上記サンプリングステップで得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割ステップと、上記帯域分割ステップで得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当てステップと、上記符号化ビット割り当てステップにおける割り当てを心理聴覚分析代替制御方式により制御するビット割り当て制御ステップと、上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化ステップと、上記量子化ステップで得られた量子化値に基づき、符号

化データを出力する符号化ステップとを実行することで、心理聴覚分析を簡略化したビット割り当て制御を実行し、人間の聴覚の特性に応じた高品質の符号化結果の得られる符号化処理を、処理負担を大きく増大することなく実行することが可能となる。

【0606】また、請求項25の音声符号化方法によれば、請求項24の方法において、上記ビット割り当て制御ステップは、上記帯域分割ステップで得られた帯域信号データに対して、心理聴覚分析代替制御方式により予め定められたビット割り当て順に従って、符号化ビット割り当てを行うよう制御する順次ビット割り当てステップであるものとしたことで、帯域ごとに単純なアルゴリズムでビットを割り当てて、簡略化した心理聴覚分析を実行し、若干の処理負担の増大はあるものの、再生音質の良好な符号化データを得ることが可能となる。

【0607】また、請求項26の音声符号化方法によれば、請求項24の方法において、上記ビット割り当て制御ステップは、上記帯域分割ステップで得られた帯域信号データに対して、心理聴覚分析代替制御方式により予め定められた各帯域への重み付けと、各帯域信号データの有する出力レベルとに基づいた符号化ビット割り当てを行うよう制御する帯域出力適応ビット割り当てステップであるものとしたことで、帯域ごとの割り当てと、データ自体の性質とに応じてビットを割り当てて、簡略化した心理聴覚分析を実行し、若干の処理負担の増大はあるものの、再生音質の良好な符号化データを得ることが可能となる。

【0608】また、請求項27の音声符号化方法によれば、請求項24の方法において、上記ビット割り当て制御ステップは、上記帯域分割ステップで得られた帯域信号データに対して、心理聴覚分析代替制御方式により予め定められた各帯域への重み付けと、各帯域毎のビット割り当て数に対する重み付けと、各帯域信号データの有する出力レベルとに基づいた符号化ビット割り当てを行うよう制御する改良型帯域出力適応ビット割り当てステップであるものとしたことで、帯域ごとの割り当てと、データ自体の性質とに応じて、すでに割り当てたビット配分を考慮しつつビットを割り当てて、簡略化した心理聴覚分析を実行し、若干の処理負担の増大はあるものの、再生音質の良好な符号化データを得ることが可能となる。

【0609】また、請求項28の音声符号化方法によれば、請求項24の方法において、上記ビット割り当て制御ステップは、上記帯域分割ステップで得られた帯域信号データに対して、帯域信号データごとに最小可聴限界値との比較を行い、上記比較により最小可聴限界未満と判定された帯域信号データにはビット割り当てを行わず、他の帯域に対してのビット割り当てを増加するよう制御する最小可聴限界比較ステップであるものとしたことで、最小可聴限界を考慮してビットを割り当てて、簡

略化した心理聴覚分析を実行し、若干の処理負担の増大はあるものの、再生音質の良好な符号化データを得ることが可能となる。

【0610】また、請求項29の映像音声符号化方法によれば、映像と音声とを符号化するにあたり、上記2つの符号化処理に含まれる処理過程の一部または全部を、共通の計算機資源を用いて実行する映像音声符号化方法において、単位時間毎の静止画像を表す複数の静止画像情報からなる原映像情報と、音声を表す原音声情報とから構成される映像音声情報が入力されたとき、上記原音声情報を一時的に蓄積する音声バッファリングステップと、上記音声バッファリングステップにおいて蓄積された原音声情報を読み出し、この読み出した上記原音声情報を符号化処理し、符号化音声情報を出力する音声符号化ステップと、映像符号化の負荷程度を表す符号化負荷基準情報を用いて、当該映像音声符号化処理についての処理能力を判断し、その判断の結果に基づいて、後述する映像符号化ステップにおける原映像情報に対する符号化を制御する符号化負荷評価ステップと、上記符号化負荷評価ステップにおける制御に従って、入力された上記原映像情報を構成する静止画像情報を符号化処理し、符号化映像情報を出力する映像符号化ステップとを実行するので、当該映像音声符号化を行う符号化装置の処理能力に対応して、映像符号化を制御し、映像音声の取り込みにもなったリアルタイムの符号化処理を実行し、再生時の音途切れのない、良好な符号化結果を得ることが可能となる。

【0611】また、請求項30の映像音声符号化方法によれば、請求項29の方法において、上記符号化負荷評価ステップは、上記原映像情報を構成する静止画像情報が入力されたとき、上記音声バッファリングステップにおいて蓄積された原音声情報の総量と、上記符号化負荷基準情報とに基づいて符号化負荷評価情報を求め、上記符号化負荷評価情報を予め設定された負荷限度と比較して、上記符号化負荷評価情報が上記負荷限度に達していない場合に静止画像情報を出力し、上記符号化負荷評価情報が上記負荷限度に達した場合に、上記静止画像情報を破棄するので、一時蓄積された音声データの量を指標として、映像データに対する符号化を実行するか、しないかを決定することで、映像音声の取り込みにもなったリアルタイムの符号化処理を実行し、再生時の音途切れのない、良好な符号化結果を得ることが可能となる。

【0612】また、請求項31の映像音声符号化方法によれば、請求項29の方法において、アナログ映像情報を入力し、後述する映像解像度情報が出力されたとき、上記アナログ映像情報を複数の離散的デジタル画素情報からなり、上記映像解像度情報に従う解像度を持つ複数の静止画像情報で構成される原映像情報に変換し、上記映像符号化ステップにおいて処理されるよう出力する映像キャプチャステップを実行するものであり、上記符号

化負荷評価ステップでは、上記音声バッファリングステップにおいて蓄積された原音声情報の総量と、映像符号化の負荷程度を表す符号化負荷基準情報とに基づいて符号化負荷評価情報を求め、上記符号化負荷評価情報に基づいて、映像符号化に用いる映像の解像度を表す映像解像度情報を求め、上記映像解像度情報を出力するものであり、上記映像符号化ステップでは、上記映像解像度情報が出力されたとき、上記映像解像度情報に従って上記静止画像情報に対して符号化処理を行い、符号化映像情報を出力するので、一時蓄積された音声データの量を指標として、映像データに対する符号化における解像度を決定することで、映像音声の取り込みにもなったリアルタイムの符号化処理を実行し、再生時の音途切れのない、良好な符号化結果を得ることが、映像のとぎれを伴わずに可能となる。

【0613】また、請求項32の映像音声符号化方法によれば、請求項29の方法において、上記符号化負荷評価ステップでは、符号化負荷評価情報を上記映像符号化ステップにおいて処理されるよう出力するものであり、上記映像符号化ステップでは、上記静止画像情報に対して、上記出力された符号化負荷評価情報を用いて計算される処理量だけ符号化処理を行い、符号化映像情報として出力するので、一時蓄積された音声データの量を指標として、映像データに対する符号化の実行割合を決定することで、映像音声の取り込みにもなったリアルタイムの符号化処理を実行し、再生時の音途切れのない、良好な符号化結果を得ることが、映像のとぎれを伴わずに可能となる。

【0614】また、請求項33の映像音声符号化方法によれば、請求項29ないし32のいずれかの方法において、上記音声符号化ステップでは、上記音声バッファリングステップにおいて蓄積された原音声情報を読み出し、この読み出した上記原音声情報の総量を計算して処理済み音声情報量として出力し、その後、上記原音声情報を符号化処理して符号化音声情報として出力するものであり、上記符号化負荷評価ステップでは、経過時間と、上記原音声情報の時間当たりの入力量に基づいて原音声入力量を求め、この原音声入力量と上記処理済み音声情報量との差である予測音声バッファ量を求め、上記予測音声バッファ量を用いて、上記符号化負荷評価情報を求めるので、一時蓄積された音声データの量の代替として予測バッファ量を指標として、映像データに対する処理を制御することで、映像音声の取り込みにもなったリアルタイムの符号化処理を実行し、再生時の音途切れのない、良好な符号化結果を得ることが、一時蓄積された音声データ量の取得が不可能、または困難な場合にも可能となる。

【0615】また、請求項34の映像音声符号化方法によれば、請求項29ないし32のいずれかの方法において、上記符号化負荷評価ステップでは、上記静止画像情

報が入力されたとき、経過時間と、上記原音声情報の時間当たりの入力量とに基づいて原音声入力量を求め、かつ、上記音声符号化ステップにおいて出力された符号化音声情報の総量に基づいて処理済み音声情報量を求め、さらに、上記求めた原音声入力量と上記求めた処理済み音声情報量との差である予測音声バッファ量を求めた後、上記予測音声バッファ量を用いて、上記符号化負荷評価情報を求めるので、一時蓄積された音声データの量の代替として予測バッファ量を指標として、映像データに対する処理を制御することで、映像音声の取り込みにともなったリアルタイムの符号化処理を実行し、再生時の音途切れのない、良好な符号化結果を得ることが、一時蓄積された音声データ量の取得と、符号化ステップからの処理量の取得とが不可能、または困難な場合にも可能となる。

【0616】また、請求項35の映像音声符号化方法によれば、請求項29ないし34のいずれかの方法において、上記符号化負荷評価ステップにおける、上記判断の結果の変動を監視し、上記変動に対応して、上記符号化負荷基準情報を設定するので、映像に対する符号化処理を制御することで、映像音声の取り込みにともなったリアルタイムの符号化処理を実行し、再生時の音途切れのない、良好な符号化結果を得ることが可能となるのに加え、制御に用いる情報の変動を監視して、当該変動に対応して、制御の基準を設定することで、再生映像画質の激しい変動を抑制することが可能となる。

【0617】また、請求項36にかかる映像符号化装置は、映像を符号化する映像符号化装置において、映像がデジタル化された、複数の静止画像情報からなる原映像情報に対して、上記静止画像情報の1つまたは複数を、後述する符号化パラメータに従って符号化する映像符号化手段と、1つ以上の解像度を一の符号化パラメータとし、フレーム内符号化、順方向予測符号化、逆方向予測符号化、及び双方向予測符号化の各タイプを含む符号化タイプのうち1つ以上の符号化タイプを他の符号化パラメータとして、上記符号化手段の処理量を決定するものである符号化パラメータを、与えられたフレームレートに基づいて決定する符号化パラメータ決定手段とを備えたので、与えられたフレームレートにおいて、高解像度の、又は高圧縮率の符号化結果の得られる符号化パラメータを設定することで、装置資源を活用して、良好な符号化結果を得ることが可能となる。

【0618】また、請求項37の音声符号化装置によれば、音声に対して、帯域分割符号化方式により符号化を行う音声符号化装置において、符号化処理に用いる数値である、設定周波数 f_s と、変換定数 n とを記憶するレジスタと、符号化の対象である音声を入力する音声入力手段と、上記記憶した設定周波数 f_s に基づいて決定されるサンプリング周波数を用いて、サンプリング音声データを作成する入力音声サンプリング手段と、上記設定

周波数 f_s をサンプリング周波数として用いた場合に得られるサンプリング音声データの個数を m 個とし、上記変換定数 n に基づいて定められる数を m' として、 m' 個のサンプリング音声データを含む、 m 個の音声データからなる変換音声データを出力する音声データ変換手段と、上記変換音声データを、帯域分割して M 個の帯域信号を得る帯域分割手段と、上記記憶した設定周波数 f_s と変換定数 n とから得られる周波数 $f_s/2n$ を制限周波数として、上記帯域信号のうち、制限周波数以下の帯域信号にのみ符号化ビットを割り当てる符号化ビット割り当て手段と、上記割り当てた符号化ビットに基づいて量子化を行う量子化手段と、上記量子化したデータを符号化データとして出力する符号化手段と、上記出力される符号化データを記録する符号化データ記録手段とを備えたので、変換定数 n の設定によって、処理負担を軽減し、音声取り込みにともなったリアルタイム符号化処理を行って、当該符号化装置の性能に応じた音質の符号化結果を得ることが可能となる。

【0619】また、請求項38の音声符号化装置によれば、音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化装置において、上記符号化に用いる制御定数を記憶する制御定数記憶手段と、入力音声をサンプリング処理して、サンプリングデータを出力するサンプリング手段と、上記サンプリング手段で得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割手段と、上記帯域分割手段で得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当て手段と、上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化手段と、上記量子化手段で得られた量子化値に基づき、符号化データを出力する符号化手段と、上記記憶した制御定数に基づいて、上記帯域分割手段、上記符号化ビット割り当て手段、上記量子化手段、および上記符号化手段におけるデータ処理を制御する符号化処理制御手段とを備えたので、制御定数の設定によって、処理負担を軽減し、音声取り込みにともなったリアルタイム符号化処理を行って、当該符号化装置の性能に応じた音質の符号化結果を得ることが可能となる。

【0620】また、請求項39の音声符号化装置によれば、音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化装置において、入力音声をサンプリング処理して、サンプリングデータを出力するサンプリング手段と、上記サンプリング手段で得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割手段と、上記帯域分割手段で得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当て手段と、上記符号化ビット割り当て手段における割り当てを心理聴覚分析代替制御方式により制御するビット割り当て制御手段と、上記符号化ビ

ットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化手段と、上記量子化手段で得られた量子化値に基づき、符号化データを出力する符号化手段とを備えたので、心理聴覚分析を簡略化したビット割り当て制御を実行し、人間の聴覚の特性に応じた高品質の符号化結果の得られる符号化処理を、処理負担を大きく増大することなく実行することが可能となる。

【0621】また、請求項40の映像音声符号化装置によれば、映像と音声とを符号化するにあたり、上記2つの符号化処理に含まれる処理過程の一部または全部を、10 共通の計算機資源を用いて実行する映像音声符号化装置において、単位時間毎の静止画像を表す複数の静止画像情報からなる原映像情報と、音声を表す原音声情報とから構成される映像音声情報が入力されたとき、上記原音声情報を一時的に蓄積する音声バッファリング手段と、上記音声バッファリング手段において蓄積された原音声情報を読み出し、この読み出した上記原音声情報を符号化処理し、符号化音声情報を出力する音声符号化手段と、映像符号化の負荷程度を表す符号化負荷基準情報を用いて、当該映像音声符号化装置の処理能力を判断し、その判断の結果に基づいて、後述する映像符号化手段に対しての上記原映像情報の出力を制御する符号化負荷評価手段と、上記符号化負荷評価手段の制御に従って、上記原映像情報を構成する静止画像情報が入力されたとき、上記静止画像情報を符号化処理し、符号化映像情報を出力する映像符号化手段とを備えたので、当該映像音声符号化を行う符号化装置の処理能力に対応して、映像符号化を制御し、映像音声の取り込みにともなったりリアルタイムの符号化処理を実行し、再生時の音途切れのない、良好な符号化結果を得ることが可能となる。 30

【0622】また、請求項41の映像符号化プログラム記録媒体によれば、映像を符号化処理する映像符号化プログラムを記録した記録媒体において、映像がデジタル化された、複数の静止画像情報からなる原映像情報に対して、上記静止画像情報の1つまたは複数、後述する符号化パラメータに従って符号化する映像符号化ステップと、1つ以上の解像度を一の符号化パラメータとし、フレーム内符号化、順方向予測符号化、逆方向予測符号化、及び双方向予測符号化の各タイプを含む符号化タイプのうち1つ以上の符号化タイプを他の符号化パラメータとして、上記符号化ステップの処理量を決定するものである符号化パラメータを、与えられたフレームレートに基づいて決定する符号化パラメータ決定ステップとを実行する符号化プログラムを記録したので、パーソナルコンピュータ等の汎用計算機において、当該符号化プログラムを実行させることにより、与えられたフレームレートにおいて、高解像度の、又は高圧縮率の符号化結果の得られる符号化パラメータを設定することで、装置資源を活用して、良好な符号化結果を得ることが可能とな 50

る。

【0623】また、請求項42の音声符号化プログラム記録媒体によれば、音声に対して、帯域分割符号化方式により符号化を行う音声符号化プログラムを記録した記録媒体において、符号化処理に用いる数値である、設定周波数 f_s と、変換定数 n とを記憶する記憶ステップと、符号化の対象である音声を入力する音声入力ステップと、上記記憶した設定周波数 f_s に基づいて決定されるサンプリング周波数を用いて、サンプリング音声データを作成する入力音声サンプリングステップと、上記設定周波数 f_s をサンプリング周波数として用いた場合に得られるサンプリング音声データの個数を m 個とし、 $m \geq m'$ である、上記変換定数 n に基づいて定められる数を m' として、 m' 個のサンプリング音声データを含む、 m 個の音声データからなる変換音声データを出力する音声データ変換ステップと、上記変換音声データを、帯域分割して M 個の帯域信号を得る帯域分割ステップと、上記記憶した設定周波数 f_s と変換定数 n とから得られる周波数 $f_s/2n$ を制限周波数として、上記帯域信号のうち、制限周波数以下の帯域信号にのみ符号化ビットを割り当てる符号化ビット割り当てステップと、上記割り当てた符号化ビットに基づいて量子化を行う量子化ステップと、上記量子化したデータを符号化データとして出力する符号化ステップと、上記出力される符号化データを記録する符号化データ記録ステップとを実行する符号化プログラムを記録したので、パーソナルコンピュータ等の汎用計算機において、当該符号化プログラムを実行させることにより、変換定数 n の設定によって、処理負担を軽減し、音声取り込みにともなったりリアルタイム符号化処理を行って、当該符号化装置の性能に応じた音質の符号化結果を得ることが可能となる。

【0624】また、請求項43の音声符号化プログラム記録媒体によれば、音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化プログラムを記録した記録媒体において、上記符号化に用いる制御定数を記憶する制御定数記憶ステップと、入力音声をサンプリング処理して、サンプリングデータを出力するサンプリングステップと、上記サンプリングステップで得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割ステップと、上記帯域分割ステップで得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当てステップと、上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化ステップと、上記量子化ステップで得られた量子化値に基づき、符号化データを出力する符号化ステップと、上記記憶した制御定数に基づいて、上記帯域分割ステップ、上記符号化ビット割り当てステップ、上記量子化ステップ、および上記符号化ステップにおけるデータ処理を制御する符号化処理制御ステップとを実行する符号化プログラム

を記録したので、パーソナルコンピュータ等の汎用計算機において、当該符号化プログラムを実行させることにより、制御定数の設定によって、処理負担を軽減し、音声取り込みにともなったりアルタイム符号化処理を行って、当該符号化装置の性能に応じた音質の符号化結果を得ることが可能となる。

【0625】また、請求項44の音声符号化プログラム記録媒体によれば、音声に対して、帯域分割符号化方式を用いて符号化を行う音声符号化プログラムを記録した記録媒体において、入力音声をサンプリング処理して、サンプリングデータを出力するサンプリングステップと、上記サンプリングステップで得られたサンプリングデータに対して帯域分割を行い、帯域信号データを出力する帯域分割ステップと、上記帯域分割ステップで得られた帯域信号データに対して、符号化ビットの割り当てを行う符号化ビット割り当てステップと、上記符号化ビット割り当てステップにおける割り当てを心理聴覚分析代替制御方式により制御するビット割り当て制御ステップと、上記符号化ビットの割り当てに従って、上記帯域信号データの量子化を行い、量子化値を出力する量子化ステップと、上記量子化ステップで得られた量子化値に基づき、符号化データを出力する符号化ステップとを実行する符号化プログラムを記録したので、パーソナルコンピュータ等の汎用計算機において、当該符号化プログラムを実行させることにより、心理聴覚分析を簡略化したビット割り当て制御を実行し、人間の聴覚の特性に応じた高品質の符号化結果の得られる符号化処理を、処理負担を大きく増大することなく実行することが可能となる。

【0626】また、請求項45の映像音声符号化プログラム記録媒体によれば、映像と音声とを符号化するにあたり、上記2つの符号化処理に含まれる処理過程の一部または全部を、共通の計算機資源を用いて実行する映像音声符号化プログラムを記録した記録媒体において、単位時間毎の静止画像を表す複数の静止画像情報からなる原映像情報と、音声を表す原音声情報とから構成される映像音声情報が入力されたとき、上記原音声情報を一時的に蓄積する音声バッファリングステップと、上記音声バッファリングステップにおいて蓄積された原音声情報を読み出し、この読み出した上記原音声情報を符号化処理し、符号化音声情報を出力する音声符号化ステップと、映像符号化の負荷程度を表す符号化負荷基準情報を用いて、当該映像音声符号化処理についての処理能力を判断し、その判断の結果に基づいて、後述する映像符号化ステップにおける原映像情報に対する符号化を制御する符号化負荷評価ステップと、上記符号化負荷評価ステップにおける制御に従って、入力された上記原映像情報を構成する静止画像情報を符号化処理し、符号化映像情報を出力する映像符号化ステップとを実行する符号化プログラムを記録したので、パーソナルコンピュータ等の

汎用計算機において、当該符号化プログラムを実行させることにより、当該映像音声符号化を行う符号化装置の処理能力に対応して、映像符号化を制御し、映像音声の取り込みにともなったりアルタイムの符号化処理を実行し、再生時の音途切れのない、良好な符号化結果を得ることが可能となる。

【図面の簡単な説明】

【図1】本発明の実施の形態1による映像符号化装置の構成を示すブロック図である。

10 【図2】同実施の形態の映像符号化装置の符号化手段における処理手順を示すフローチャート図である。

【図3】同実施の形態の映像符号化装置の符号化パラメータ決定手段における処理手順を示すフローチャート図である。

【図4】本発明の実施の形態2による映像符号化装置の構成を示すブロック図である。

【図5】同実施の形態の映像符号化装置の符号化パラメータ決定手段における処理手順を示すフローチャート図である。

20 【図6】本発明の実施の形態3による映像符号化装置の構成を示すブロック図である。

【図7】同実施の形態の映像符号化装置の符号化パターン決定手段における状態遷移を示す状態遷移図である。

【図8】本発明の実施の形態4による映像符号化装置の構成を示すブロック図である。

【図9】同実施の形態の映像符号化装置の符号化パターン決定手段における状態遷移を示す状態遷移図である。

【図10】本発明の実施の形態5による音声符号化装置の構成を示すブロック図である。

30 【図11】同実施の形態の装置のハードウェア構成を示す図である。

【図12】同実施の形態の装置による音声符号化の処理手順を示すフローチャート図である。

【図13】同実施の形態の装置によるサンプリング処理と音声データ変換処理を説明するための図である。

【図14】本発明の実施の形態6による音声符号化装置の構成を示すブロック図である。

【図15】同実施の形態の装置による音声符号化の処理手順を示すフローチャート図である。

40 【図16】本発明の実施の形態7による音声符号化装置の構成を示すブロック図である。

【図17】同実施の形態の装置による音声符号化の処理手順を示すフローチャート図である。

【図18】本発明の実施の形態8による音声符号化装置の構成を示すブロック図である。

【図19】同実施の形態の装置による音声符号化の処理手順を示すフローチャート図である。

【図20】本発明の実施の形態9による音声符号化装置の構成を示すブロック図である。

50 【図21】同実施の形態の装置による音声符号化の処理

手順を示すフローチャート図である。

【図 2 2】同実施の形態による音声符号化を説明するための概念図である。

【図 2 3】本発明の実施の形態 10 による音声符号化装置の構成を示すブロック図である。

【図 2 4】同実施の形態の装置による音声符号化の処理手順を示すフローチャート図である。

【図 2 5】同実施の形態による帯域分割での基本低域通過フィルタの演算式における係数 C_i を示す図である。

【図 2 6】本発明の実施の形態 11 による音声符号化装置の構成を示すブロック図である。

【図 2 7】同実施の形態の装置による音声符号化の処理手順を示すフローチャート図である。

【図 2 8】同実施の形態による音声符号化における帯域の間引きを説明するための概念図である。

【図 2 9】本発明の実施の形態 12 による音声符号化装置の構成を示すブロック図である。

【図 3 0】同実施の形態の装置による音声符号化の処理手順を示すフローチャート図である。

【図 3 1】本発明の実施の形態 13 による音声符号化装置の構成を示すブロック図である。

【図 3 2】同実施の形態の装置による音声符号化の処理手順を示すフローチャート図である。

【図 3 3】本発明の実施の形態 14 による音声符号化装置の構成を示すブロック図である。

【図 3 4】同実施の形態の装置による音声符号化の処理手順を示すフローチャート図である。

【図 3 5】同実施の形態の装置による順次ビット割り当ての処理手順を示すフローチャート図である。

【図 3 6】同実施の形態による順次ビット割り当てにおける、各帯域へのビット割り当ての処理手順を示すフローチャート図である。

【図 3 7】本発明の実施の形態 15 による音声符号化装置の構成を示すブロック図である。

【図 3 8】同実施の形態の装置による音声符号化の処理手順を示すフローチャート図である。

【図 3 9】本発明の実施の形態 16 による音声符号化装置の構成を示すブロック図である。

【図 4 0】同実施の形態の装置による音声符号化の処理手順を示すフローチャート図である。

【図 4 1】同実施の形態の装置による改良型帯域出力適応ビット割り当ての処理手順を示すフローチャート図である。

【図 4 2】本発明の実施の形態 17 による音声符号化装置の構成を示すブロック図である。

【図 4 3】同実施の形態の装置による音声符号化の処理手順を示すフローチャート図である。

【図 4 4】本発明の実施の形態 18 による映像音声符号化装置の概略構成を示す図である。

【図 4 5】同実施の形態の映像音声符号化装置の動作を

図解的に表した図である。

【図 4 6】同実施の形態の映像音声符号化装置のより長期の時間にわたる動作を説明するための図である。

【図 4 7】本発明の実施の形態 19 による映像音声符号化装置の概略構成を示す図である。

【図 4 8】本発明の実施の形態 20 による映像音声符号化装置の概略構成を示す図である。

【図 4 9】本発明の実施の形態 21 による映像音声符号化装置の概略構成を示す図である。

【図 5 0】本発明の実施の形態 22 による映像音声符号化装置の概略構成を示す図である。

【図 5 1】実施の形態 18 ないし実施の形態 22 による映像音声符号化における現象を説明するための図である。

【図 5 2】本発明の実施の形態 23 による映像音声符号化装置の概略構成を示す図である。

【図 5 3】実施の形態 23 による映像音声符号化を説明するための図である。

【図 5 4】本発明の実施の形態 24 による映像音声符号化装置の概略構成を示す図である。

【図 5 5】本発明の実施の形態 25 による映像符号化装置の構成を示すブロック図である。

【図 5 6】同実施の形態の映像符号化装置の符号化手段における処理手順を示すフローチャート図である。

【図 5 7】同実施の形態の映像符号化装置の符号化パラメータ決定手段における処理手順を示すフローチャート図である。

【図 5 8】従来の技術による専用のハードウェアで構成されたリアルタイム処理を行う映像符号化装置の構成を示すブロック図である。

【図 5 9】従来技術の第 1 例による音声符号化装置の構成を示すブロック図である。

【図 6 0】同例における音声符号化の処理手順を示すフローチャート図である。

【図 6 1】音声符号化におけるサンプリング処理を説明するための図である。

【図 6 2】音声符号化における帯域分割を説明するための概念図である。

【図 6 3】帯域分割された帯域信号を示す図である。

【図 6 4】従来技術の第 2 例による MPEG 1 Audio 音声符号化装置の構成を示すブロック図である。

【図 6 5】同例における音声符号化処理を示すフローチャート図である。

【図 6 6】従来技術による心理聴覚分析に応用される、人間の聴覚における最小可聴限界を示す図である。

【図 6 7】従来の技術による映像音声符号化装置の概略構成を示す図である。

【符号の説明】

101, 201, 301, 401, 5001 符号化

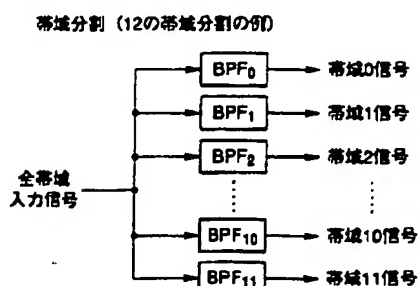
手段

159

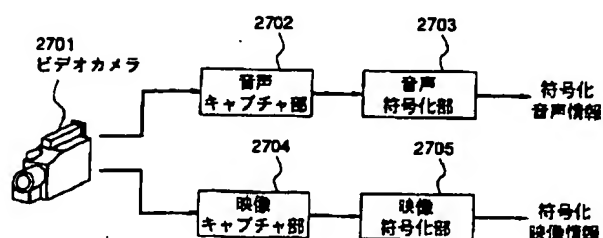
160

102, 202, 302, 402, 5002	符号化	* 1606, 1706, 2556, 2656符号化ビット
パラメータ決定手段		割り当て部
103, 203, 303, 403, 5003	DCT	507, 607, 707, 807, 907, 1007,
処理手段		1107, 1207, 1307, 1407, 1507,
104, 204, 304, 404, 5004	量子化	1607, 1707, 2557, 2657量子化部
手段		508, 608, 708, 808, 908, 1008,
105, 205, 305, 405, 5005	可変長	1108, 1208, 1308, 1408, 1508,
符号化手段		1608, 1708, 2558, 2658符号化部
106, 206, 306, 406, 5006	ビット	509, 609, 709, 809, 909, 1009,
ストリーム生成手段	10	1109, 1209, 1309, 1409, 1509,
107, 207, 307, 407, 5007	逆量子	1609, 1709, 2559, 2659符号化データ
化		記録部
108, 208, 308, 408, 5008	逆DC	7010, 1213 入力バッファ
T処理手段		7011, 1214 入力バッファ監視部
109, 209, 309, 409, 5009	予測画	8012, 1315 符号化データ監視部
像生成手段		1718 最小可聴限界比較部
110 解像度参照テーブル		2660 FFT (高速フーリエ変換) 部
210 符号化パターン参照テーブル		2661 心理聴覚分析部
310, 410 符号化パターン決定手段		1801, 1901, 2001, 2101, 2201,
311, 411 処理能力判断手段	20	2301, 2401, 2701ビデオカメラ
412 バッファ手段		1802, 1902, 2002, 2102, 2202,
413 入力フレームレート制御手段		2302, 2402, 2702音声キャプチャ部
501, 601, 701, 801, 901, 1001,		1803, 1903, 2003, 2103, 2203,
1101, 1201, 1301, 1401, 1501,		2303, 2403音声バッファリング部
1601, 1701, 2551, 2651音声入力部		1805, 1905, 2005, 2142, 2205,
502, 602, 702, 802, 902, 1011,		2305, 2405, 2703音声符号化部
1112, 1202, 1302レジスタ		1806, 1906, 2006, 2106, 2206,
810, 1210, 1310 固定的符号化レジスタ		2306, 2406, 2704映像キャプチャ部
503, 603, 703, 803, 903, 1003,		1807, 1923, 2035, 2107, 2207,
1103, 1203, 1303, 1403, 1503,	30	2307, 2407, 2705映像符号化部
1603, 1703, 2553, 2653入力音声サン		1808, 1921, 2032, 2144, 2253,
プリング部		2308, 2408符号化負荷評価部
504, 604, 704, 804 音声データ変換部		1924 フレーム間予測処理部
1118 帯域間引き部		1925 フレーム符号化部
505, 605, 705, 805, 905, 1005,		2037 解像度補正情報付加部
1105, 1205, 1305, 1405, 1505,		2141, 2251, 2361, 2461 システム
1605, 1705, 2555, 2655帯域分割部		タイマ
506, 606, 706, 806, 906, 1006,		2411 符号化負荷提示部
1106, 1206, 1306, 1406, 1506, *		2412 負荷設定用標準映像音声出力部

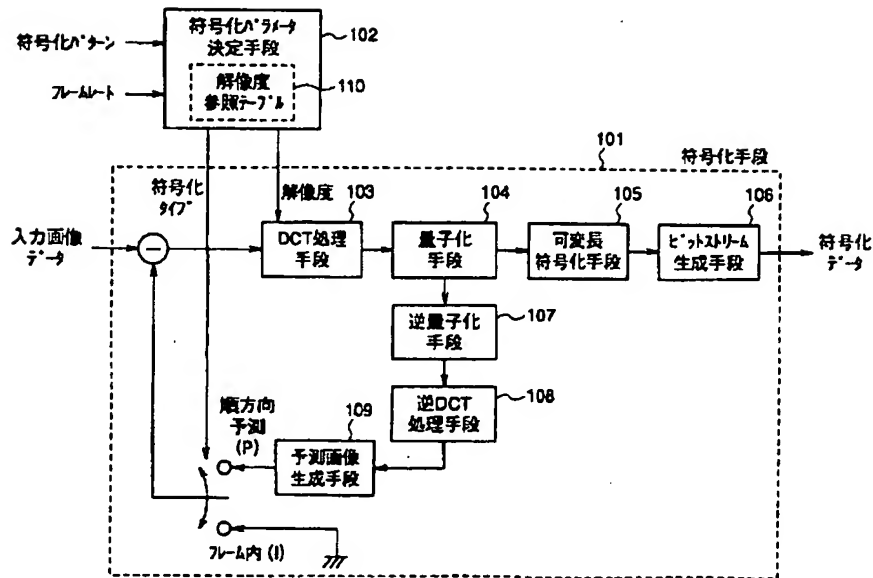
【図62】



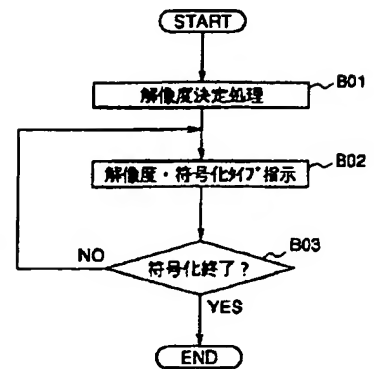
【図67】



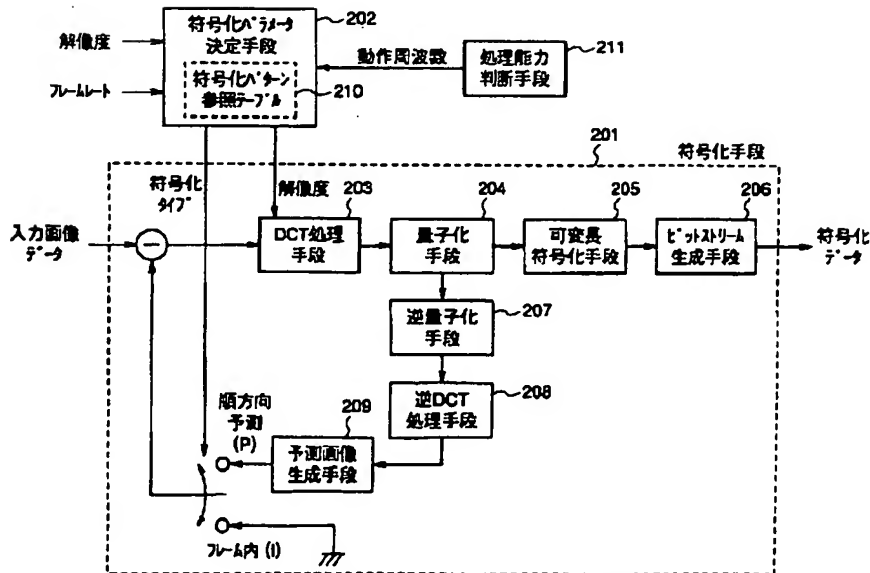
【図1】



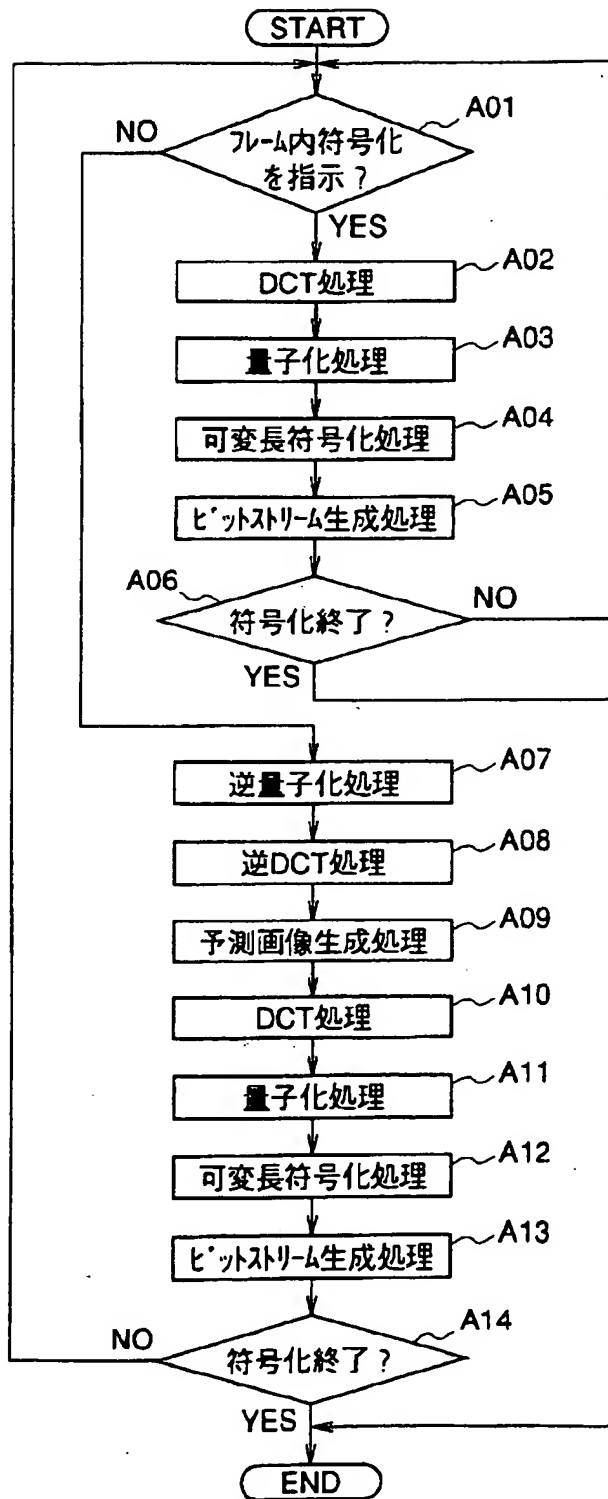
【図3】



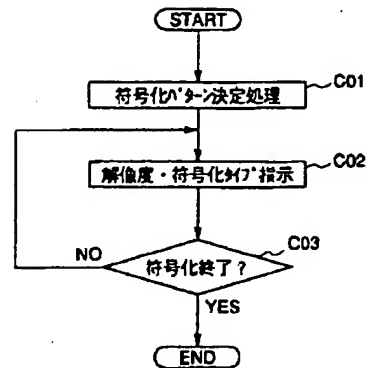
【図4】



【図2】

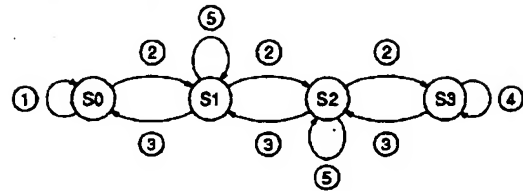


【図5】



【図7】

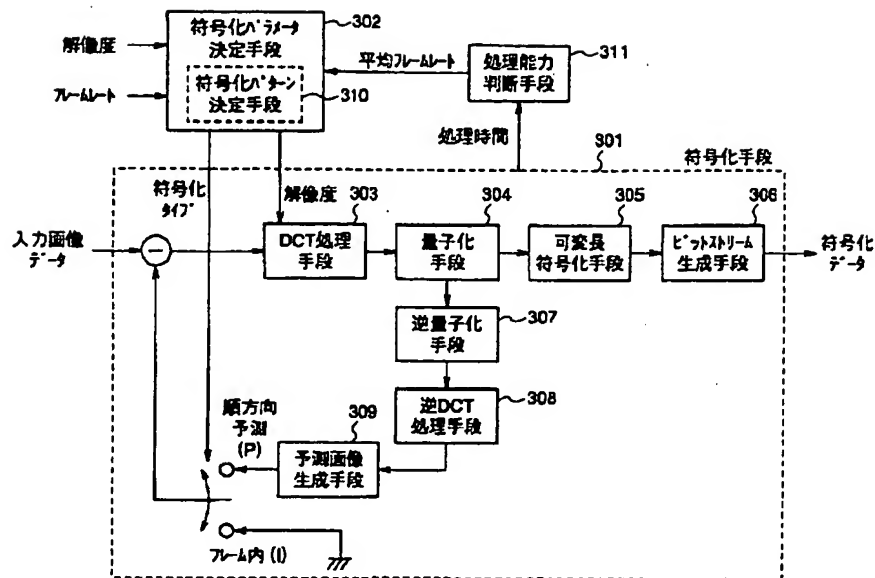
(a) 状態遷移図



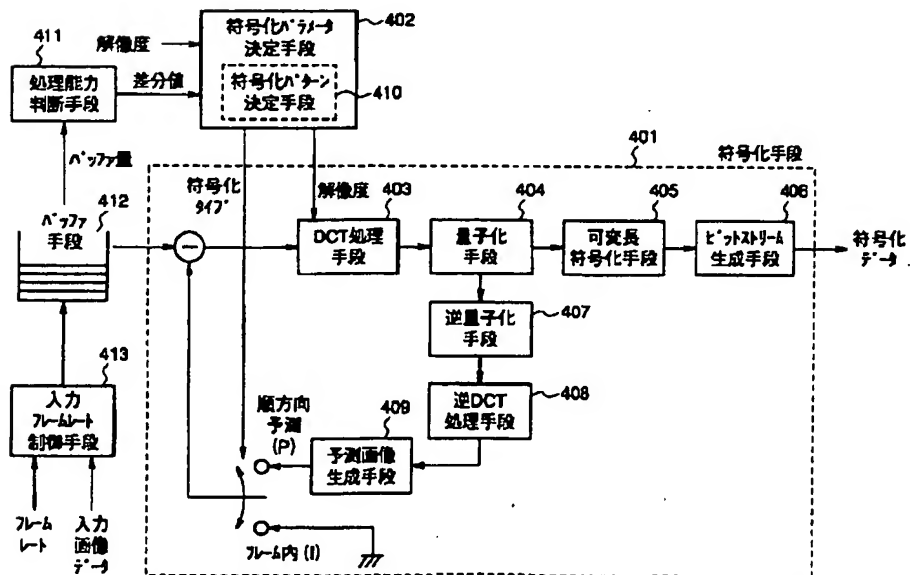
(b) 状態遷移条件

遷移条件 ①	② 以外のとき
遷移条件 ②	指定されたフレームレート < 処理能力判断手段より通知されるフレームレート
遷移条件 ③	指定されたフレームレート > 処理能力判断手段より通知されるフレームレート
遷移条件 ④	③ 以外のとき
遷移条件 ⑤	指定されたフレームレート = 処理能力判断手段より通知されるフレームレート

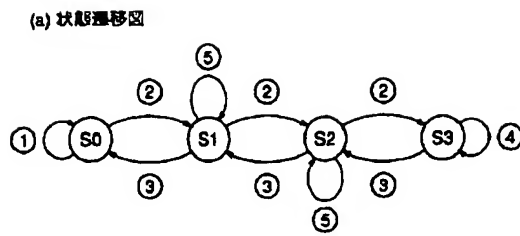
【図6】



【図8】



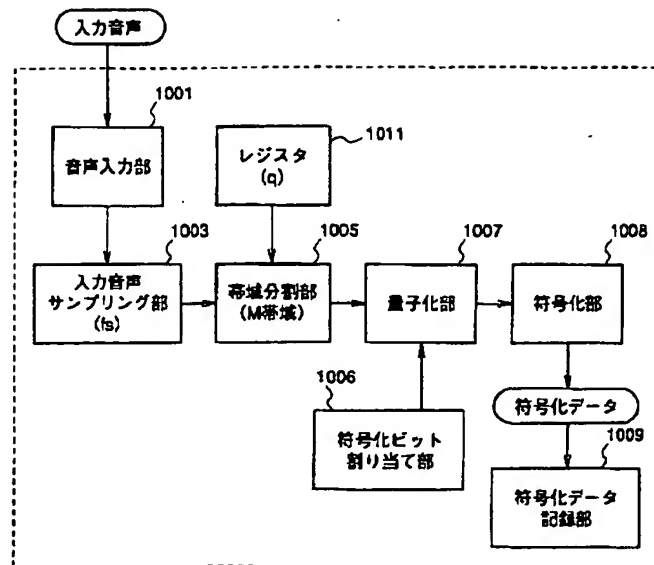
【図9】



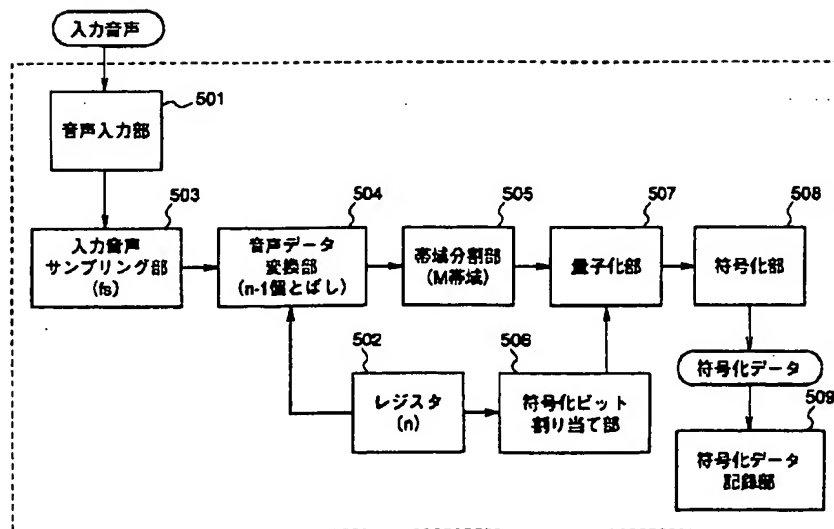
(b) 状態遷移条件

遷移条件 ①	② 以外するとき
遷移条件 ②	処理能力判断手段より通知される 差分値が負の値のとき
遷移条件 ③	処理能力判断手段より通知される 差分値が正の値のとき
遷移条件 ④	③ 以外するとき
遷移条件 ⑤	処理能力判断手段より通知される 差分値が0のとき

【図23】

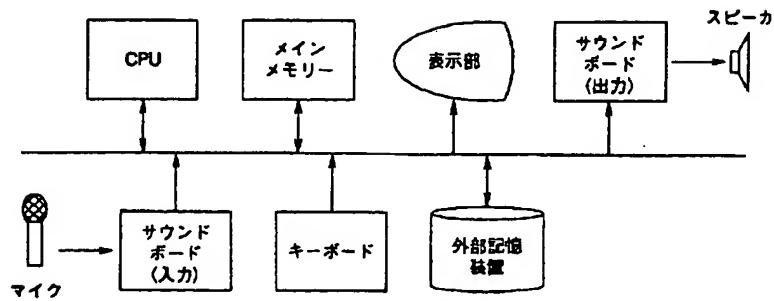


【図10】

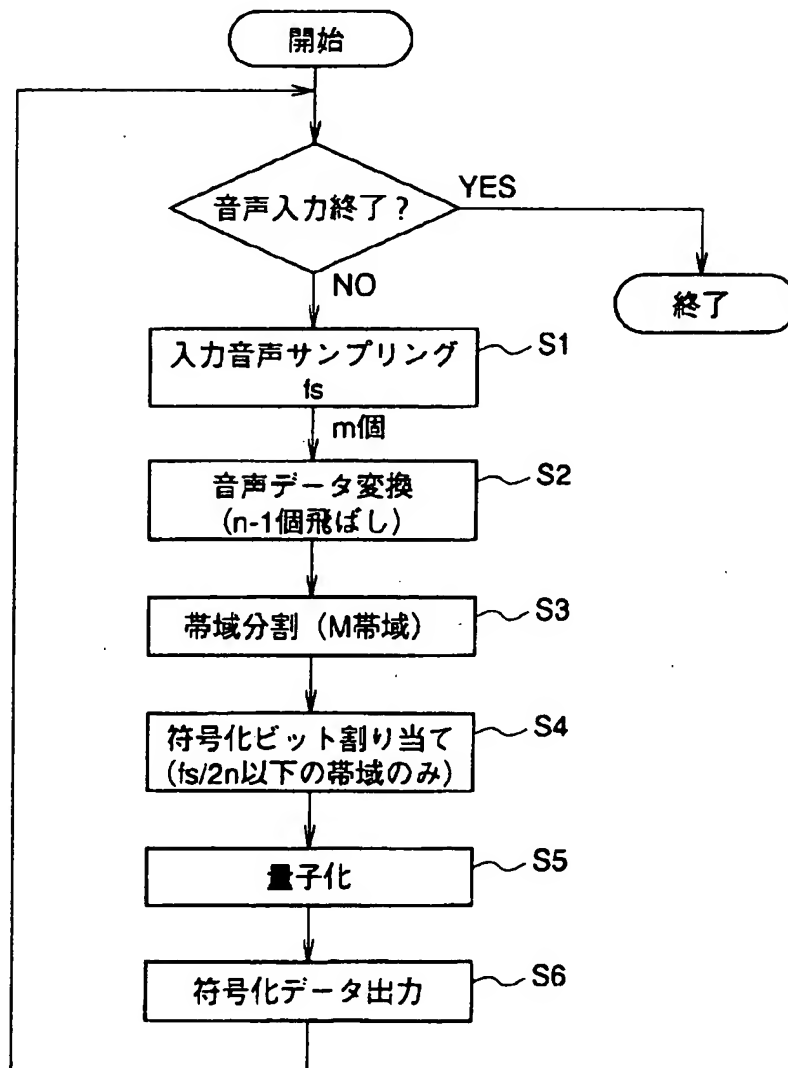


【図11】

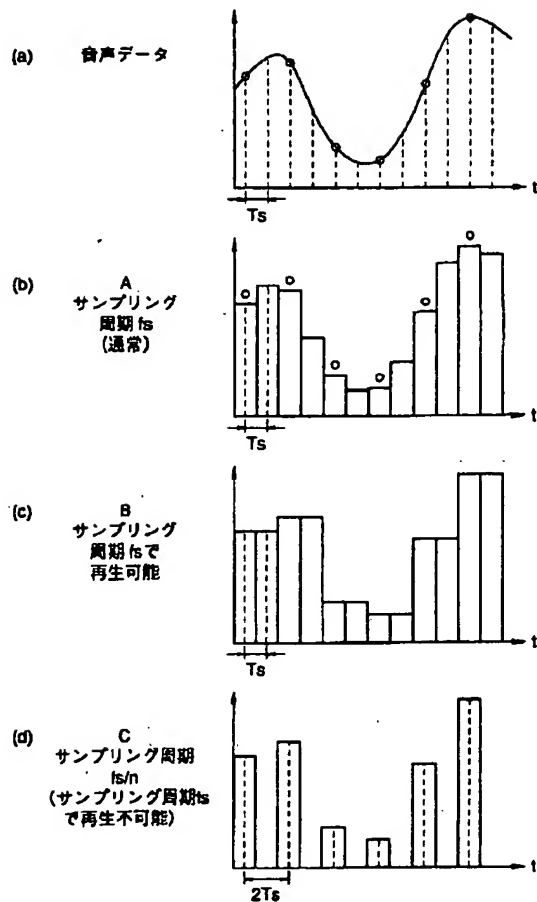
音声符号化装置



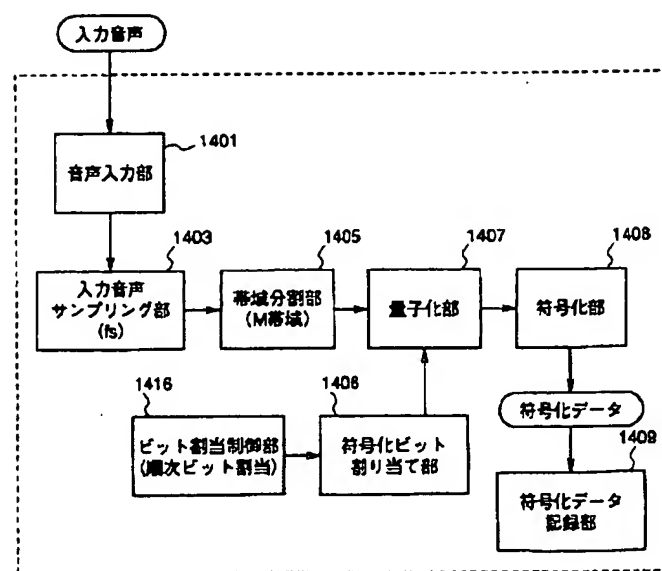
【図12】



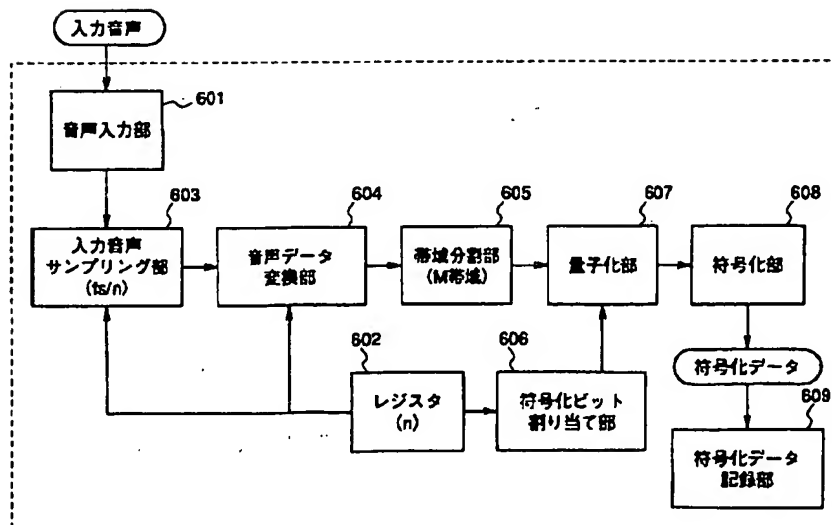
【図13】



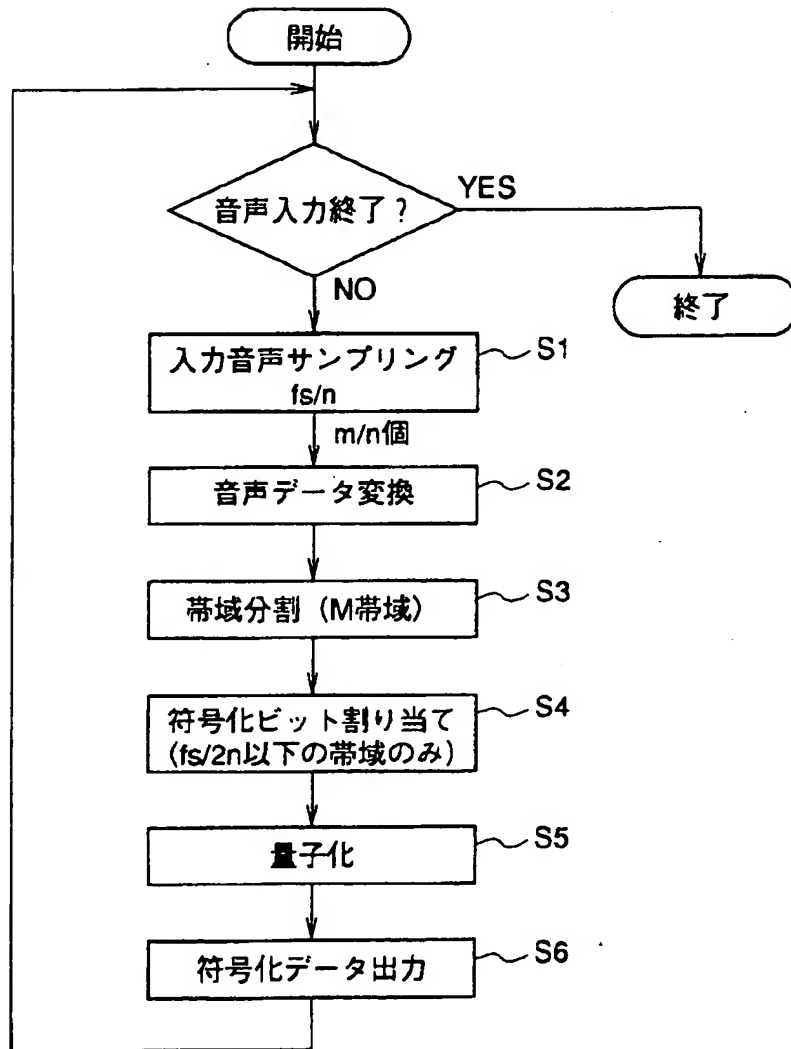
【図33】



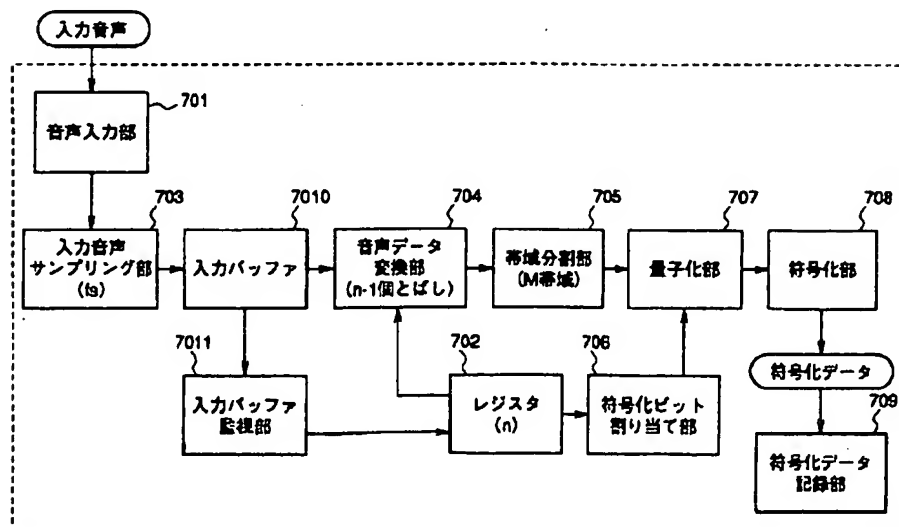
【図14】



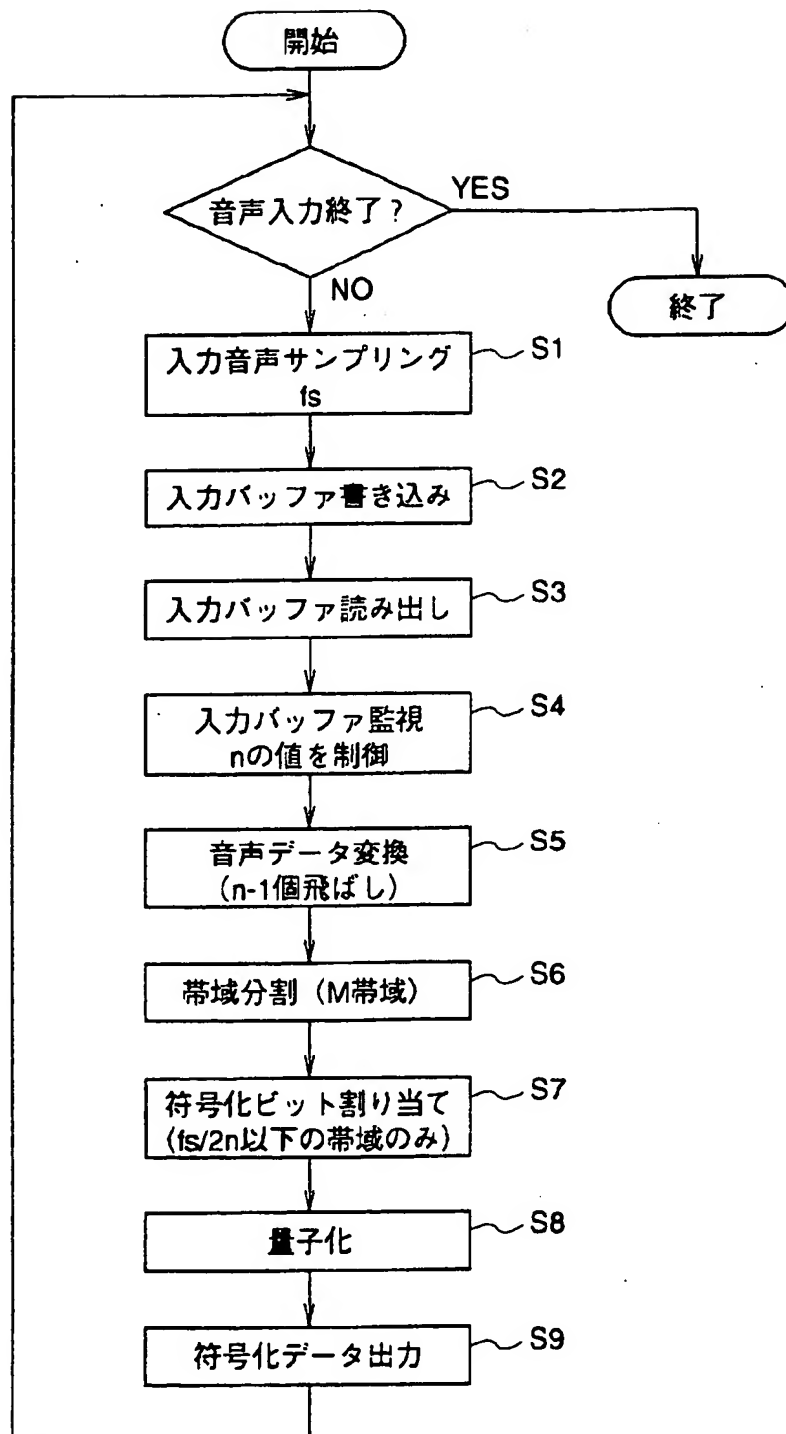
【図15】



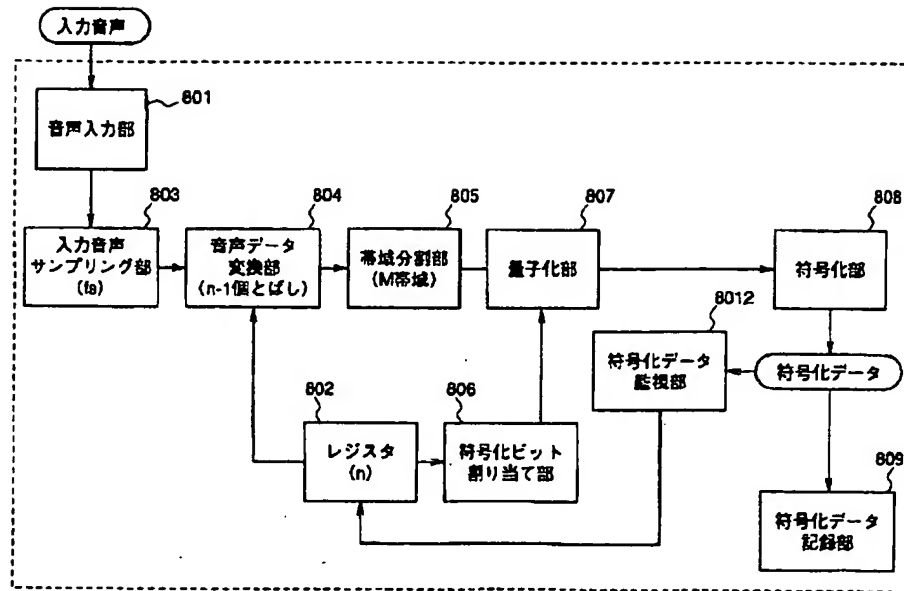
【図16】



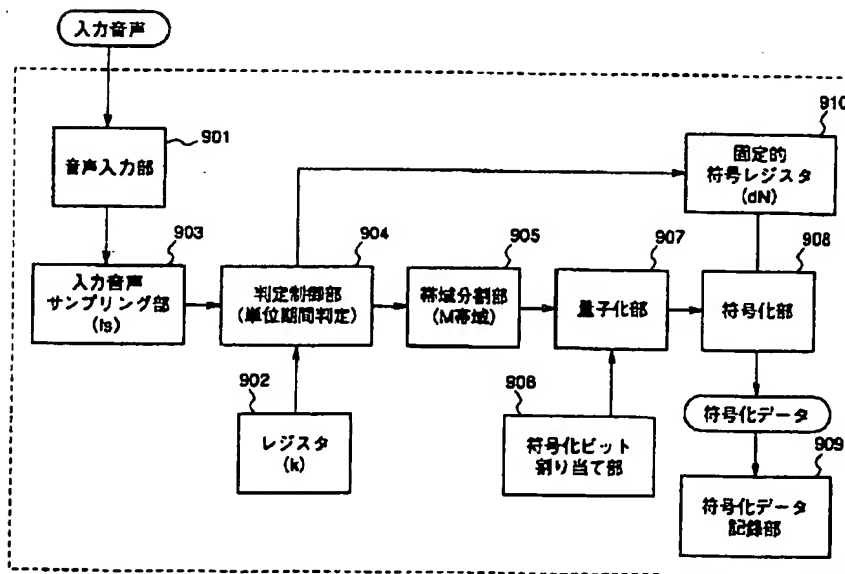
【図17】



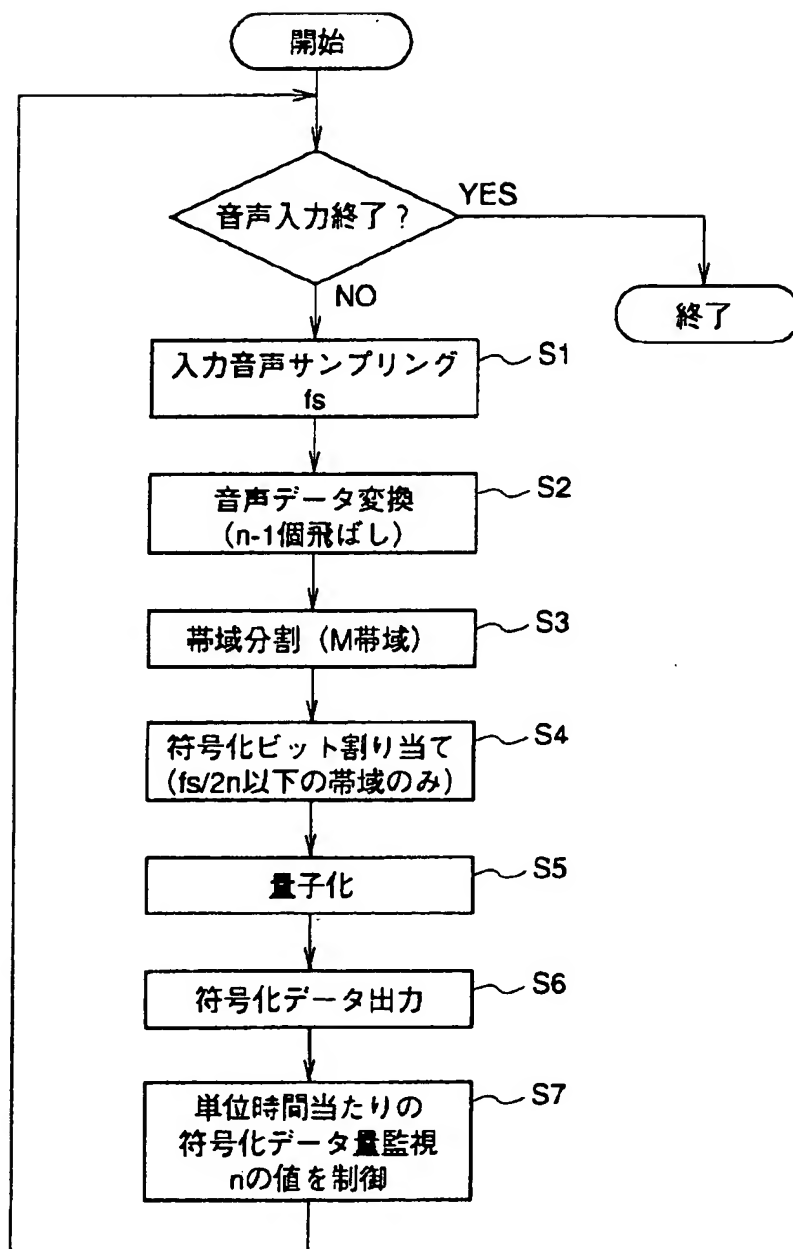
【図18】



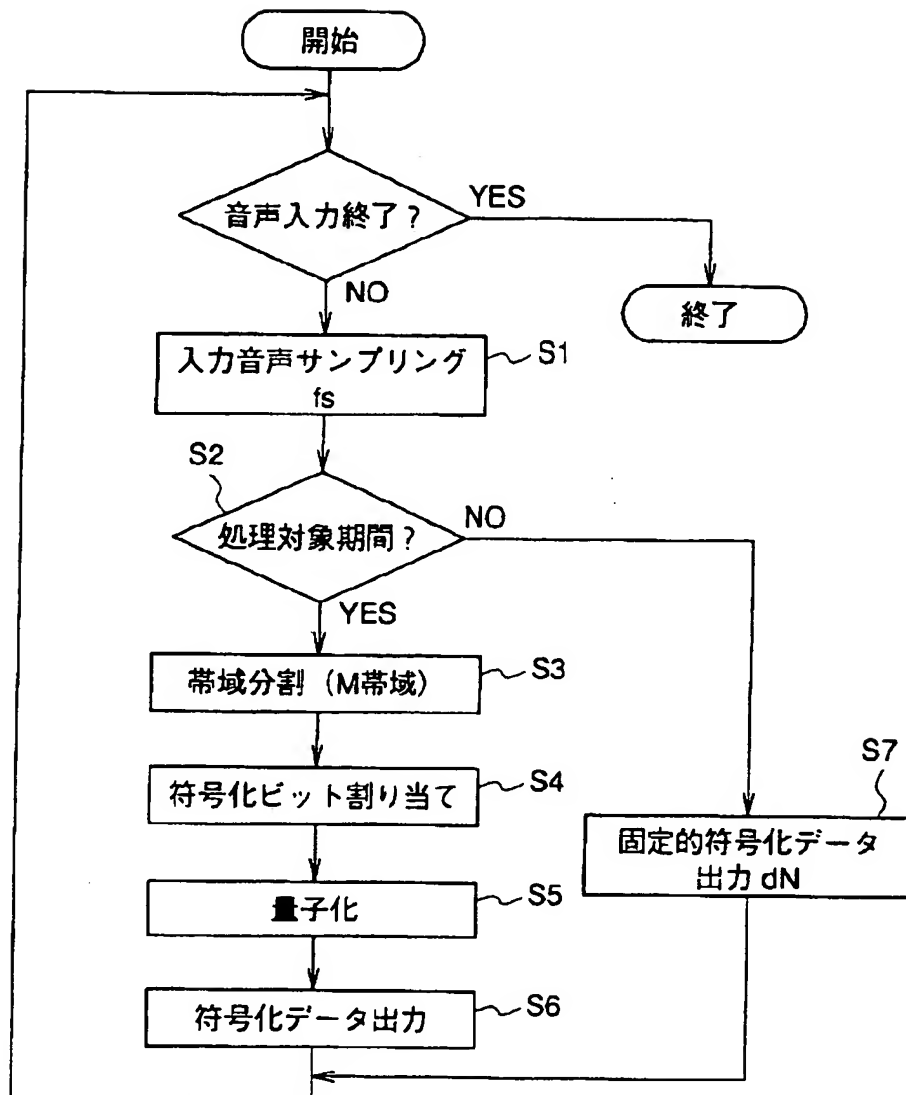
【図20】



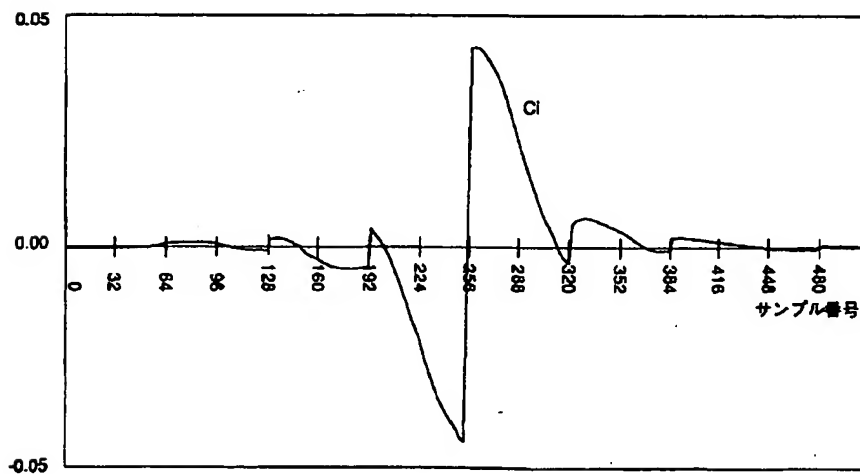
【図19】



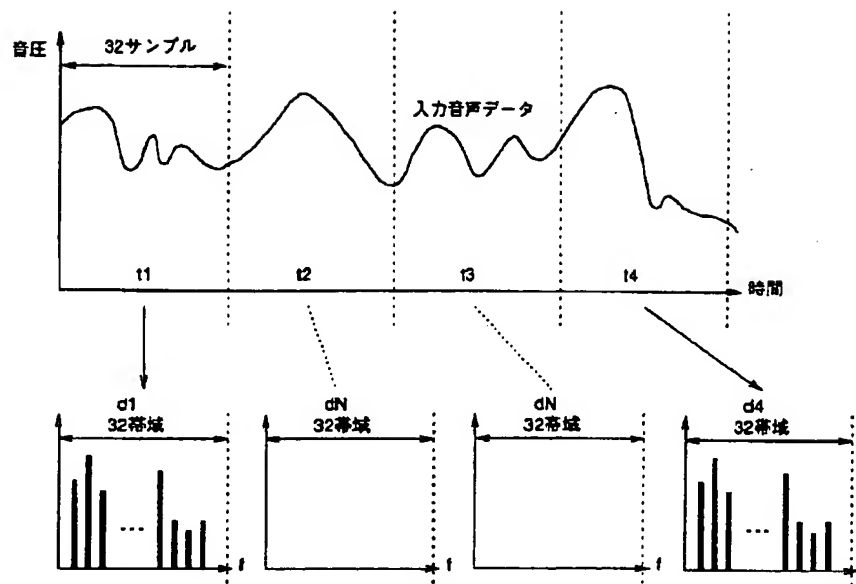
【図21】



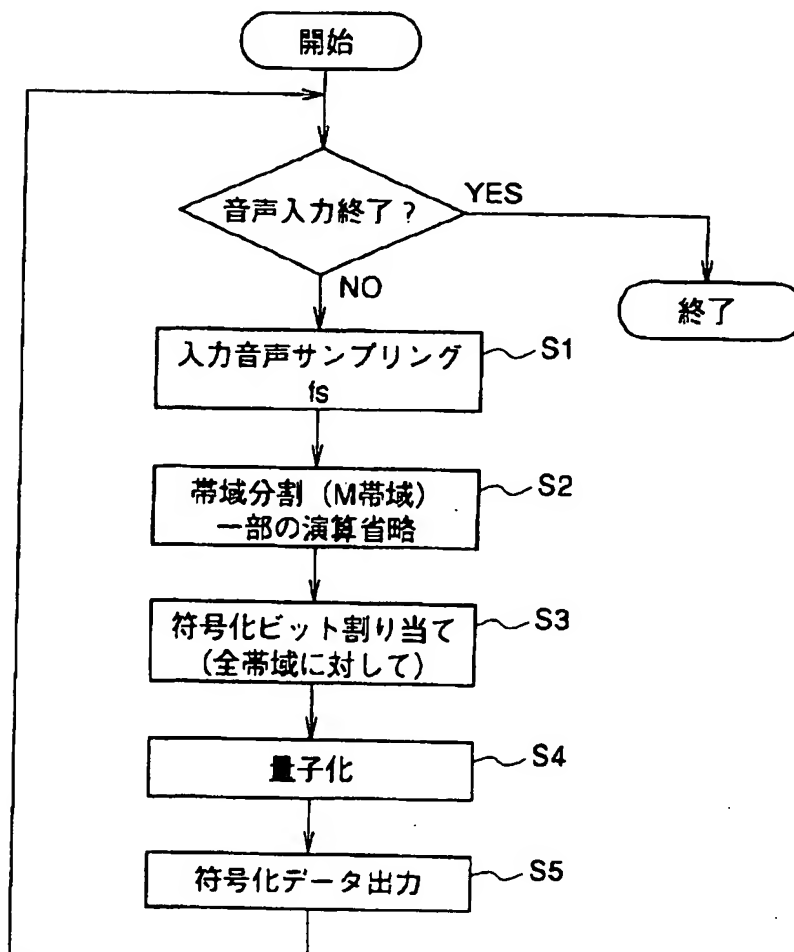
【図25】



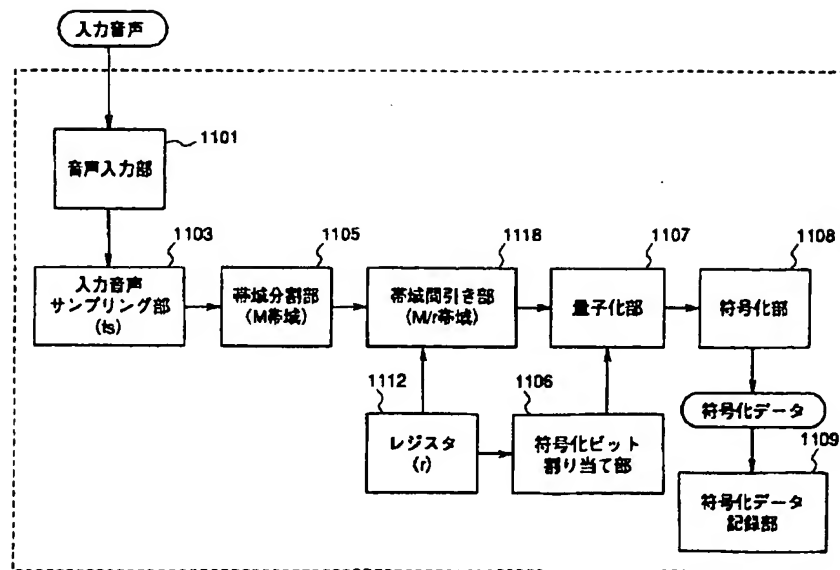
【図22】



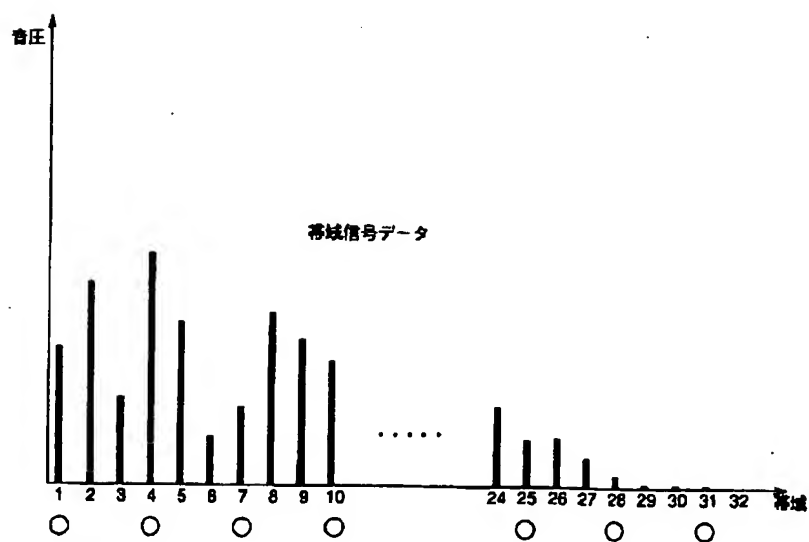
【図24】



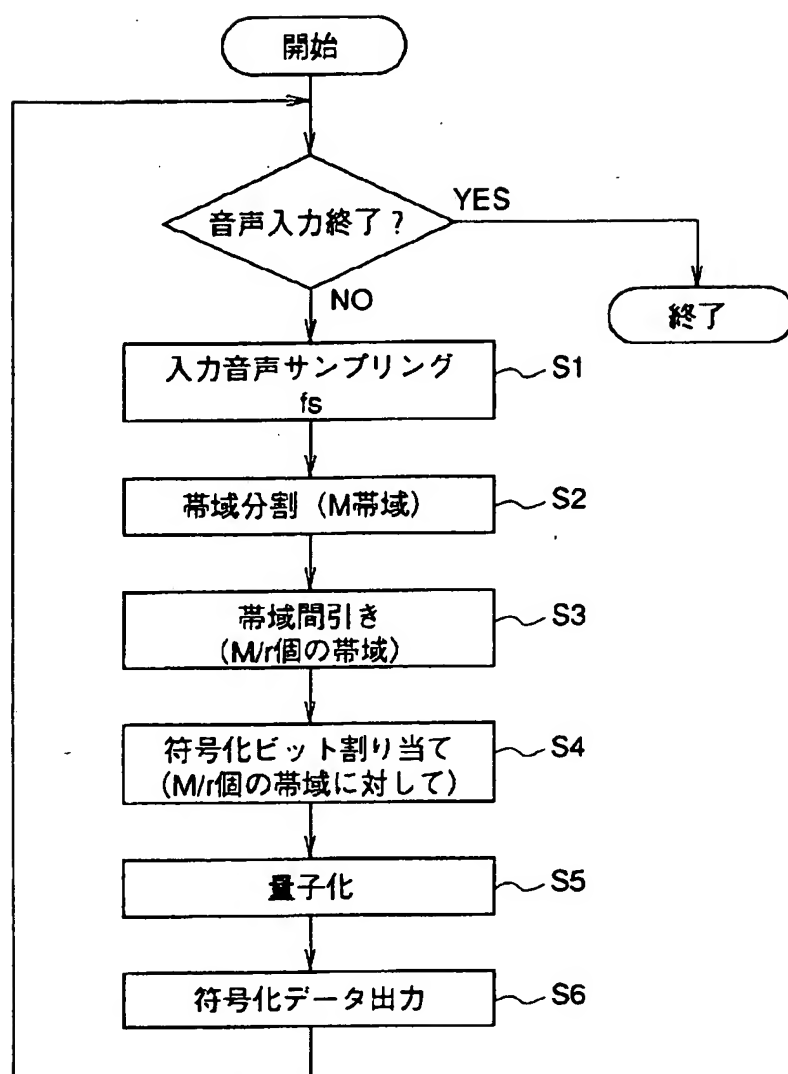
【図26】



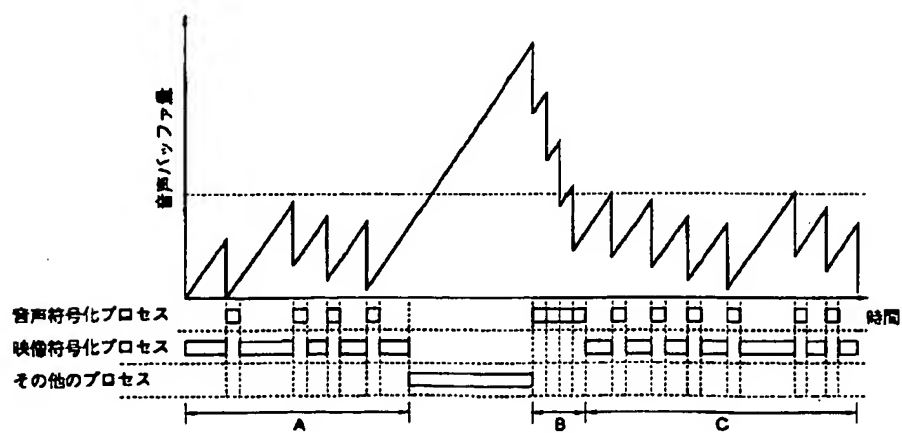
【図28】



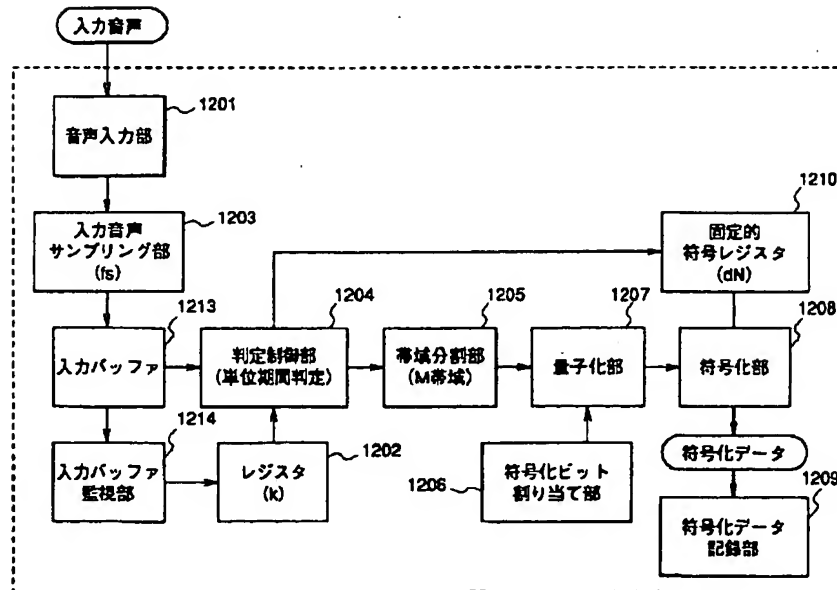
【図27】



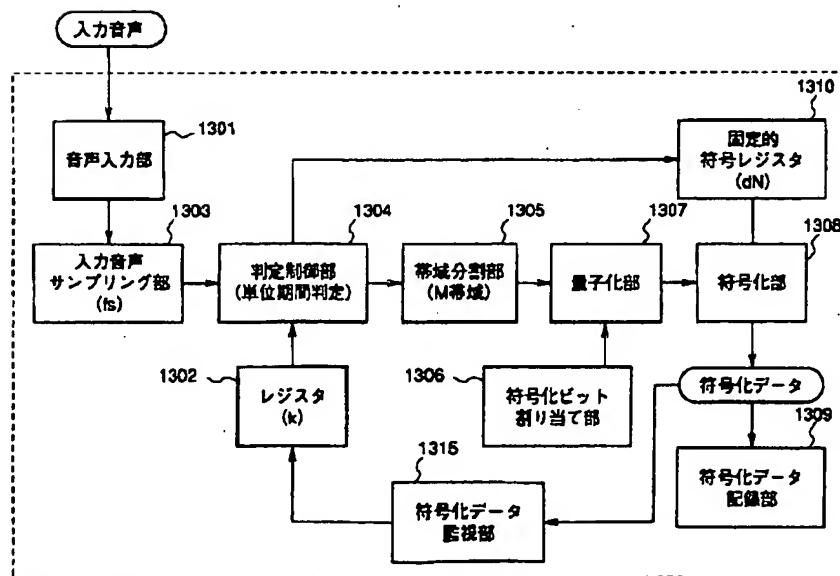
【図46】



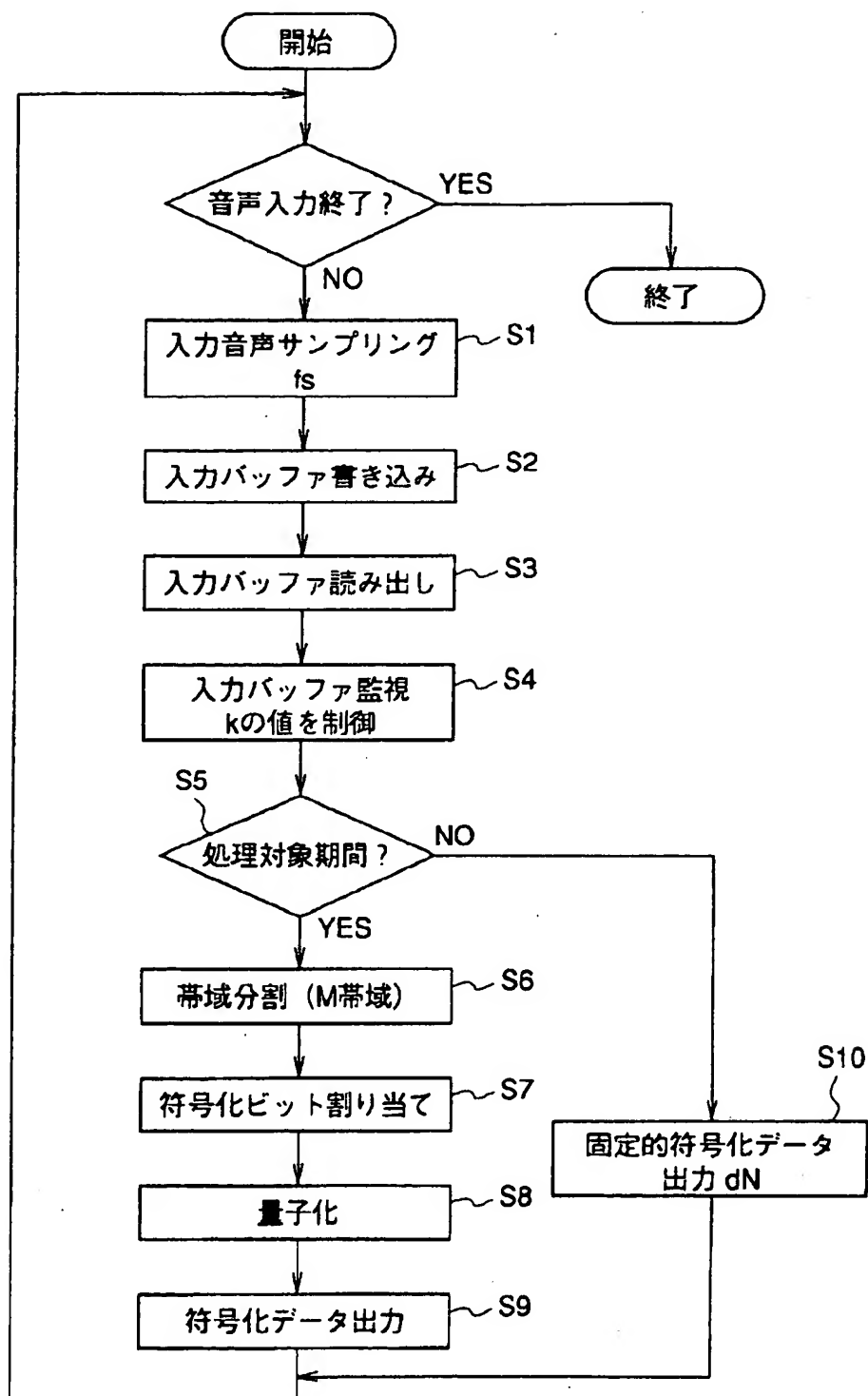
【図29】



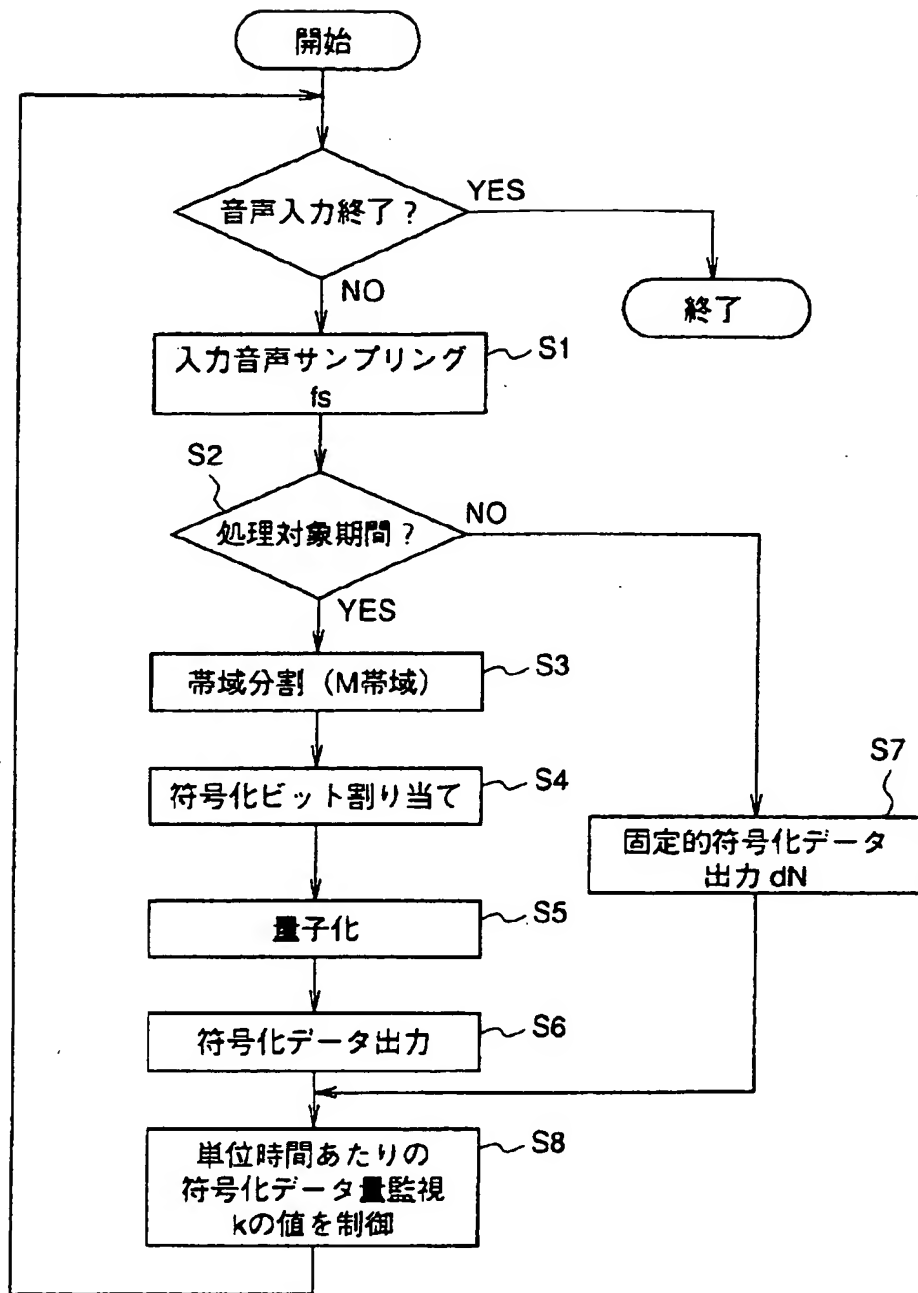
【図31】



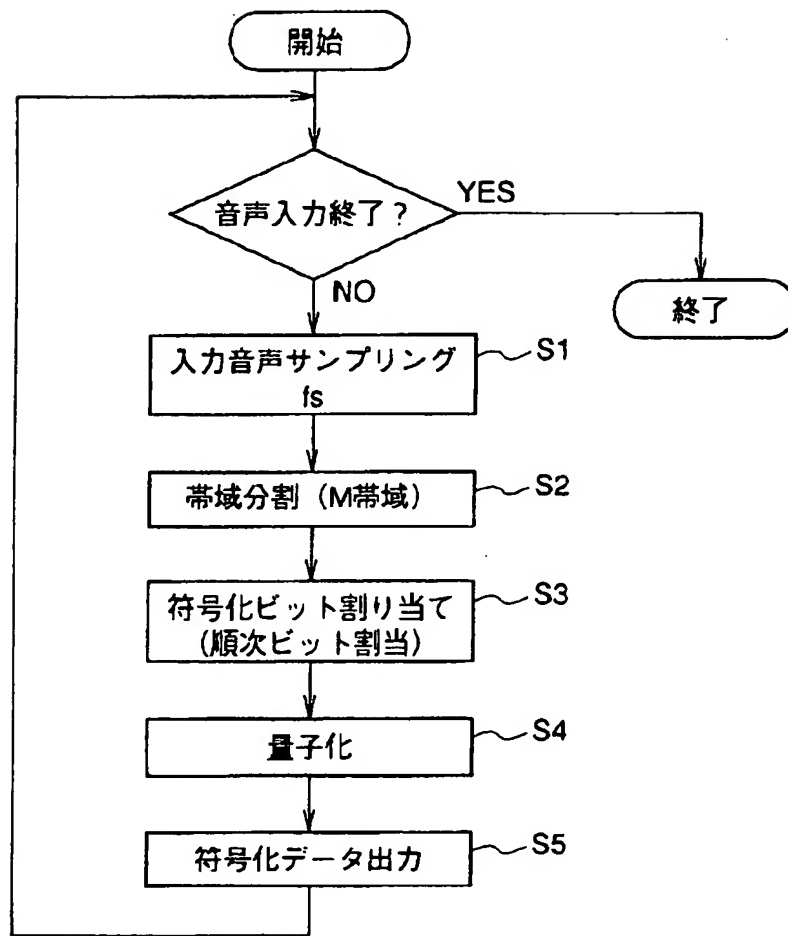
【図30】



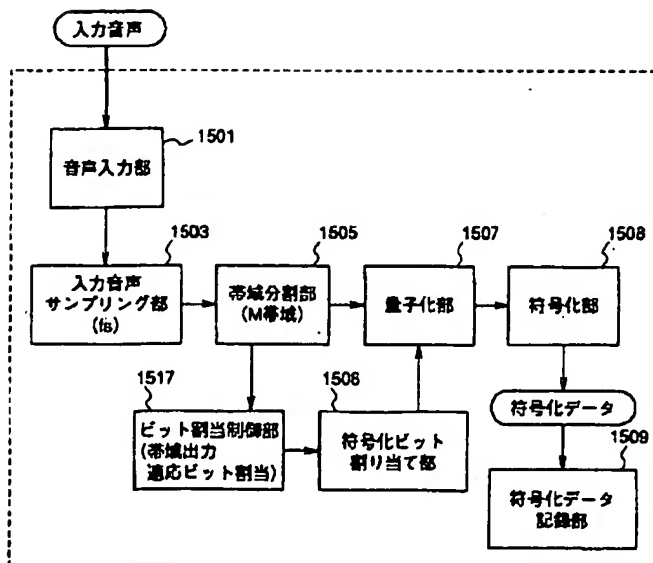
【図32】



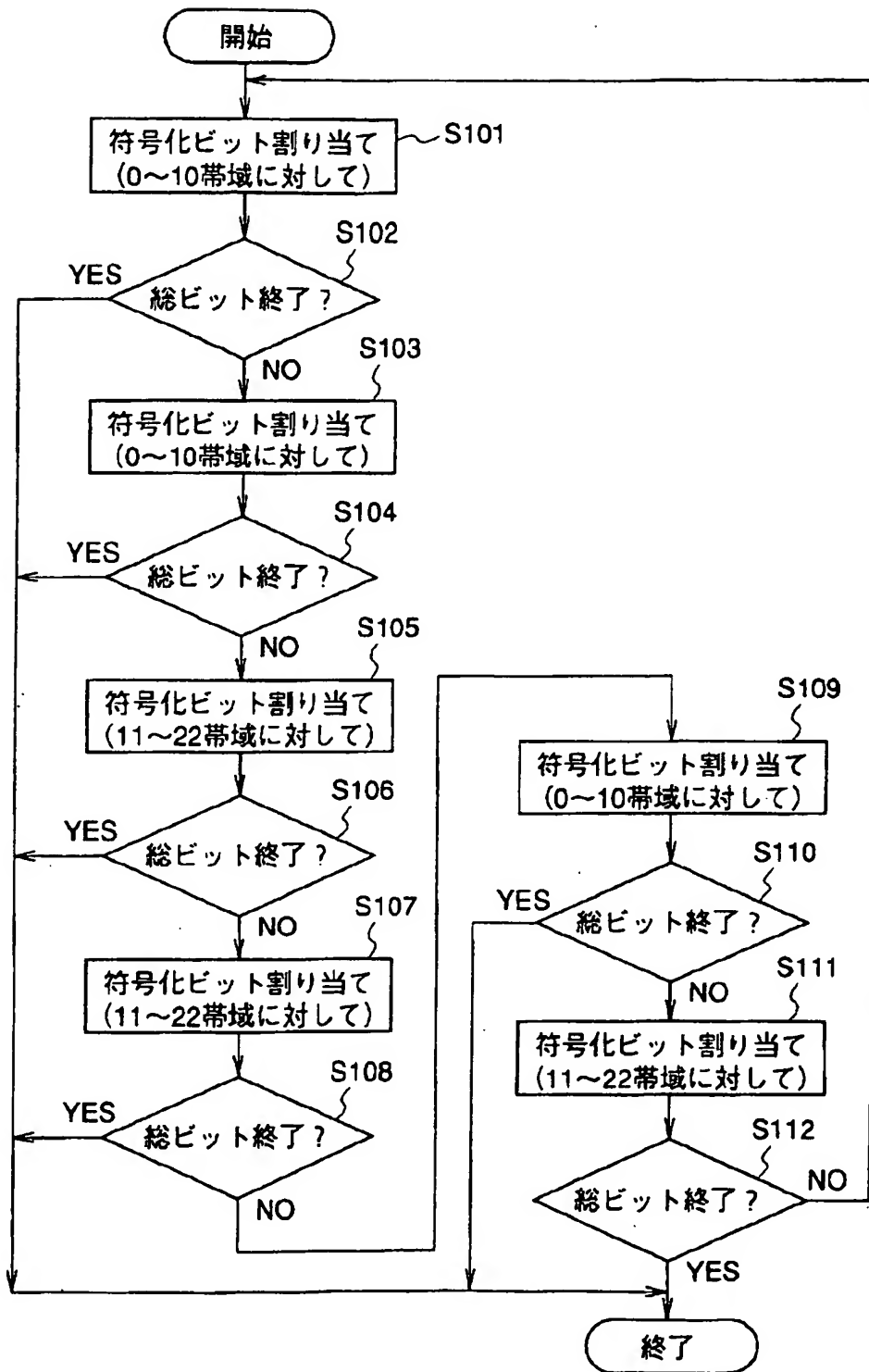
【図34】



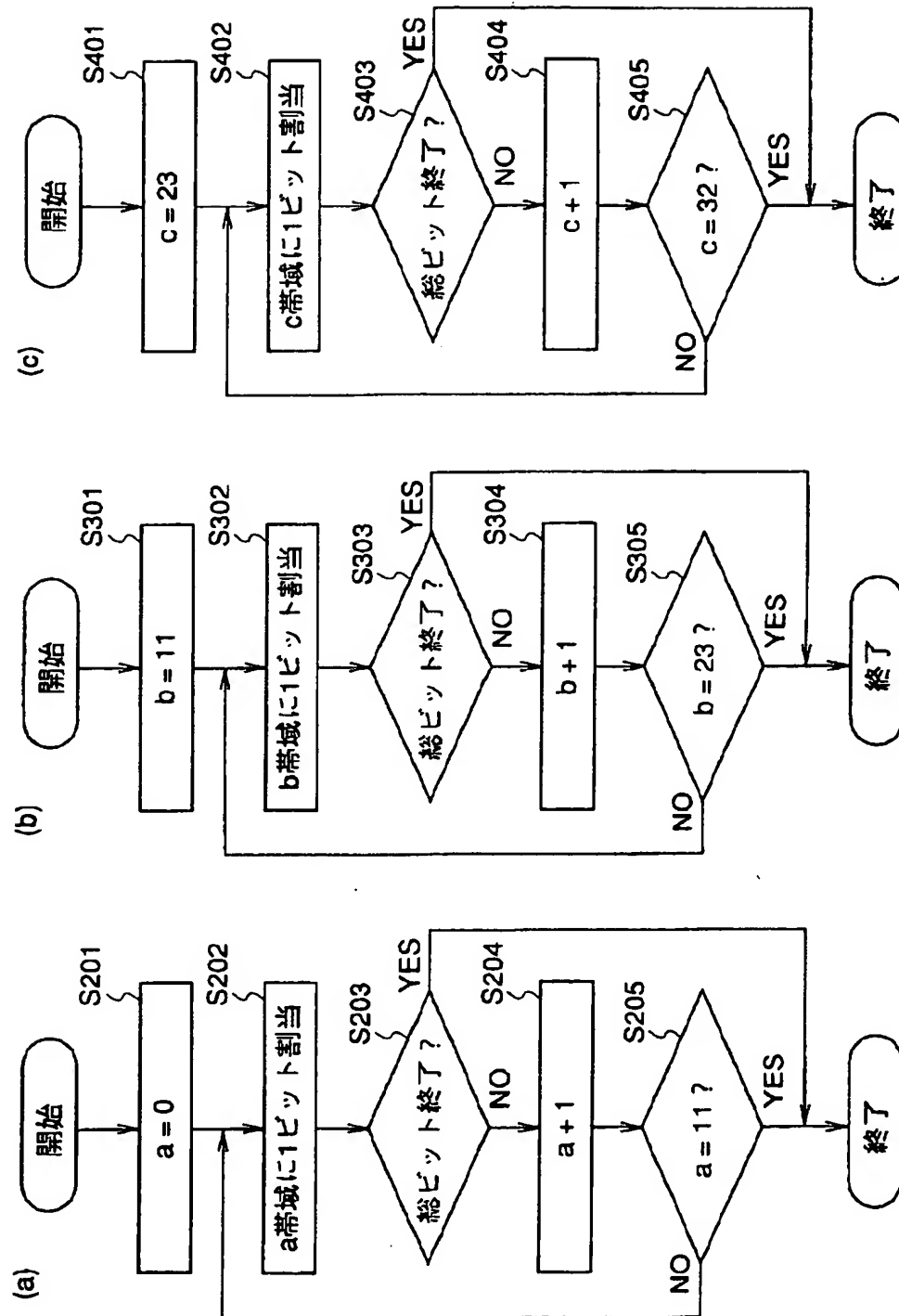
【図37】



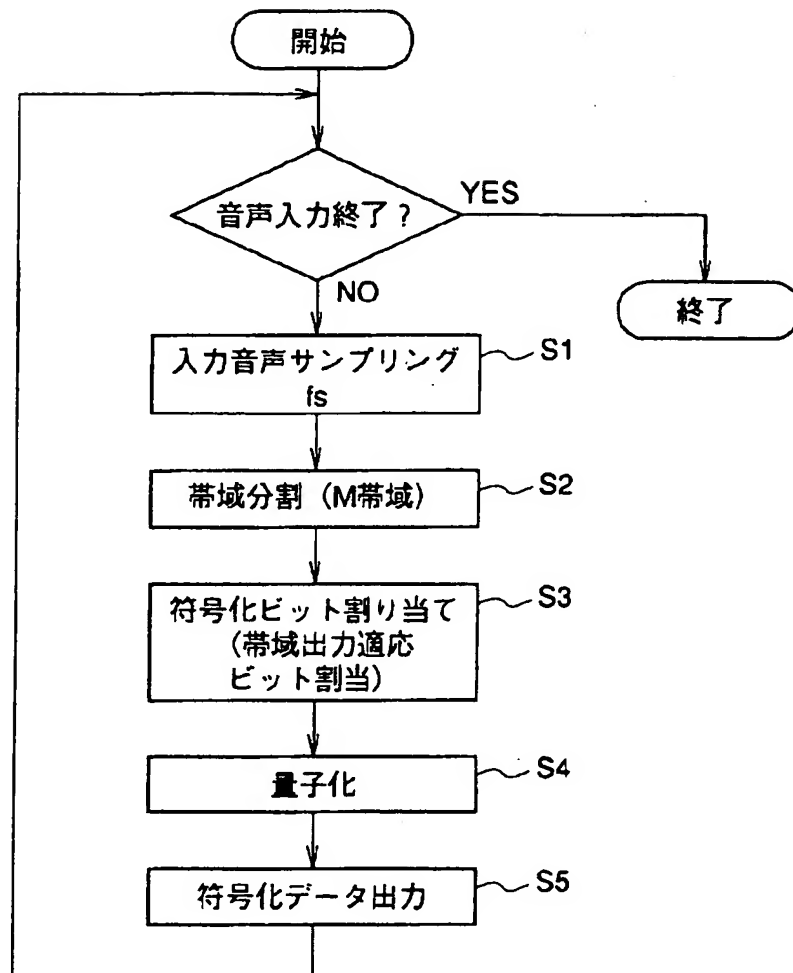
【図35】



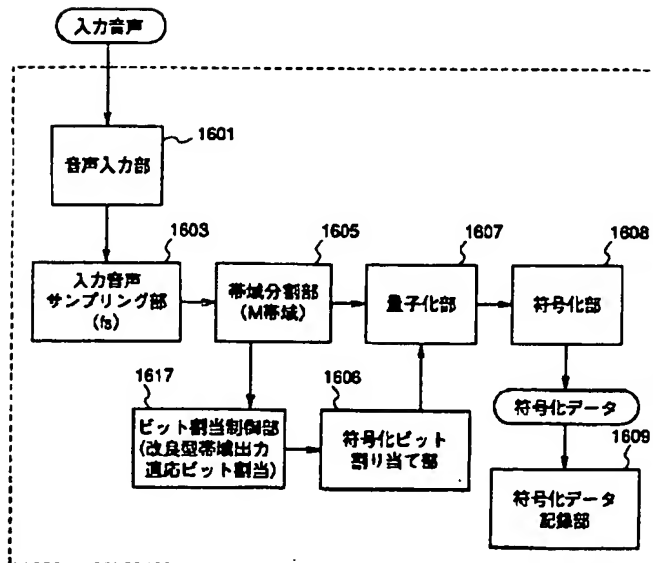
【図36】



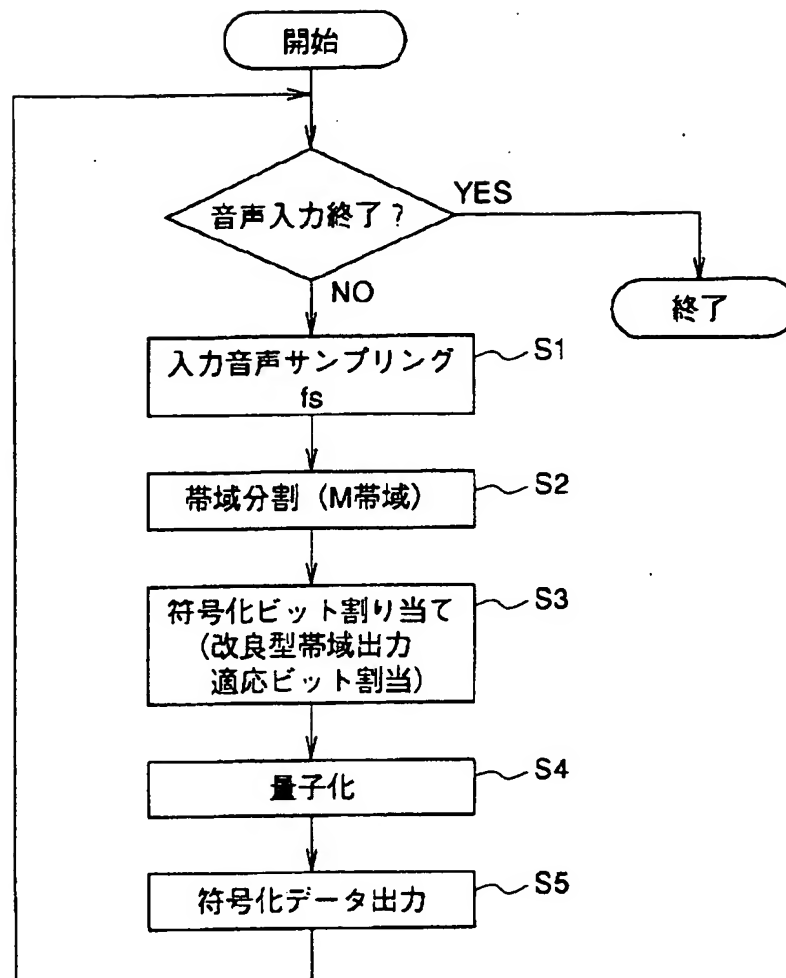
【図38】



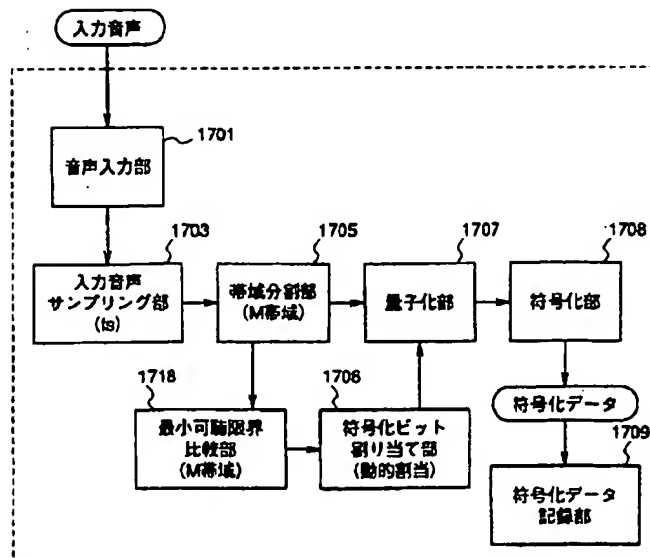
【図39】



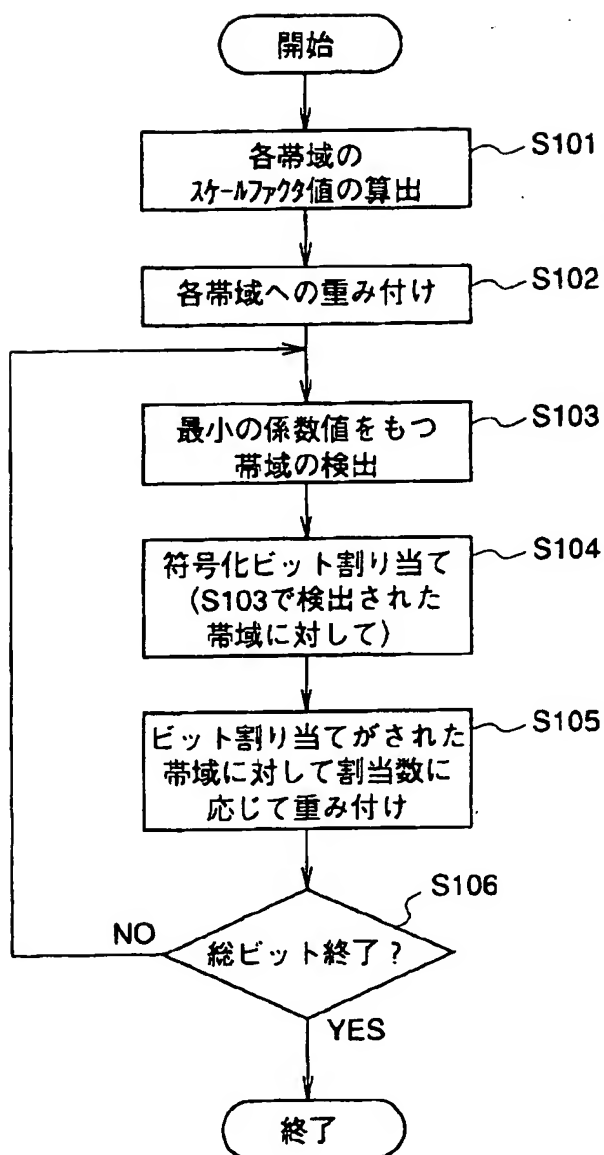
【図40】



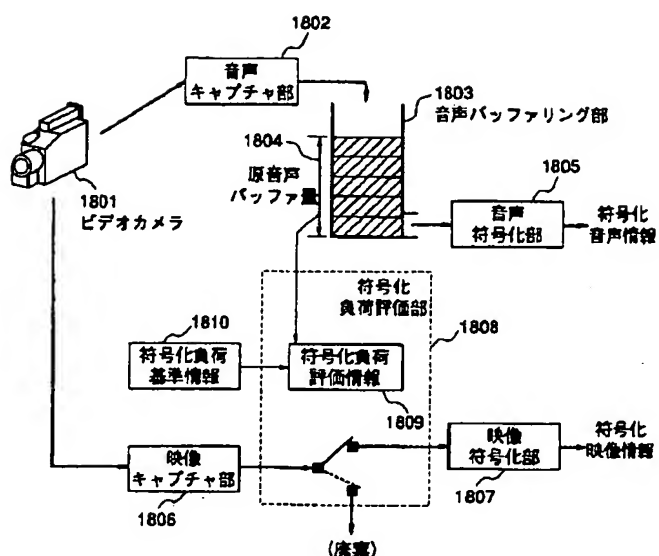
【図42】



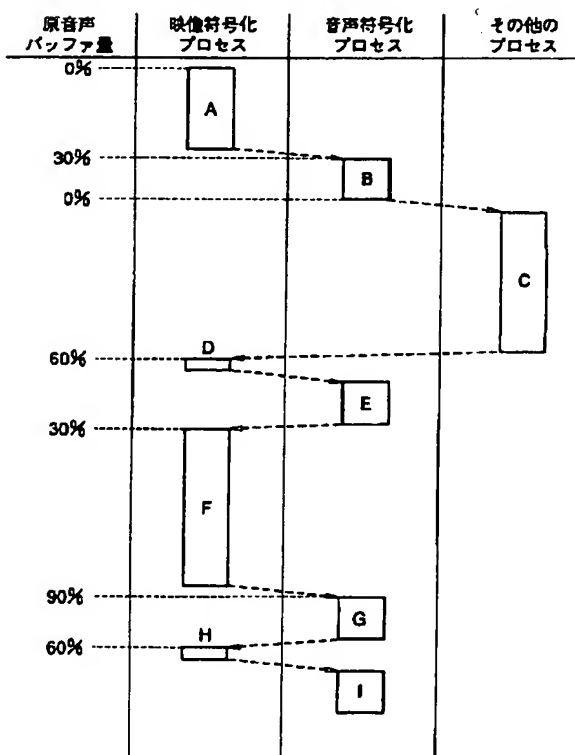
【図41】



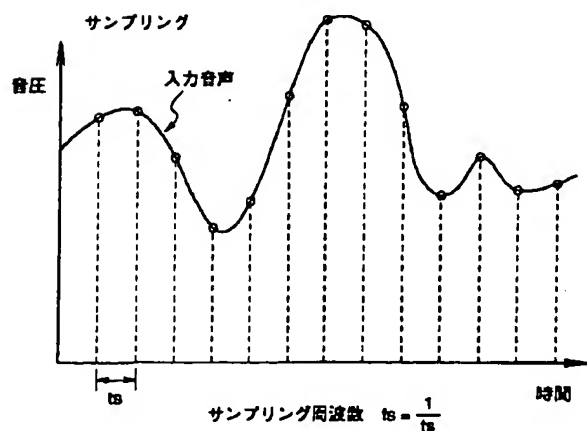
【図44】



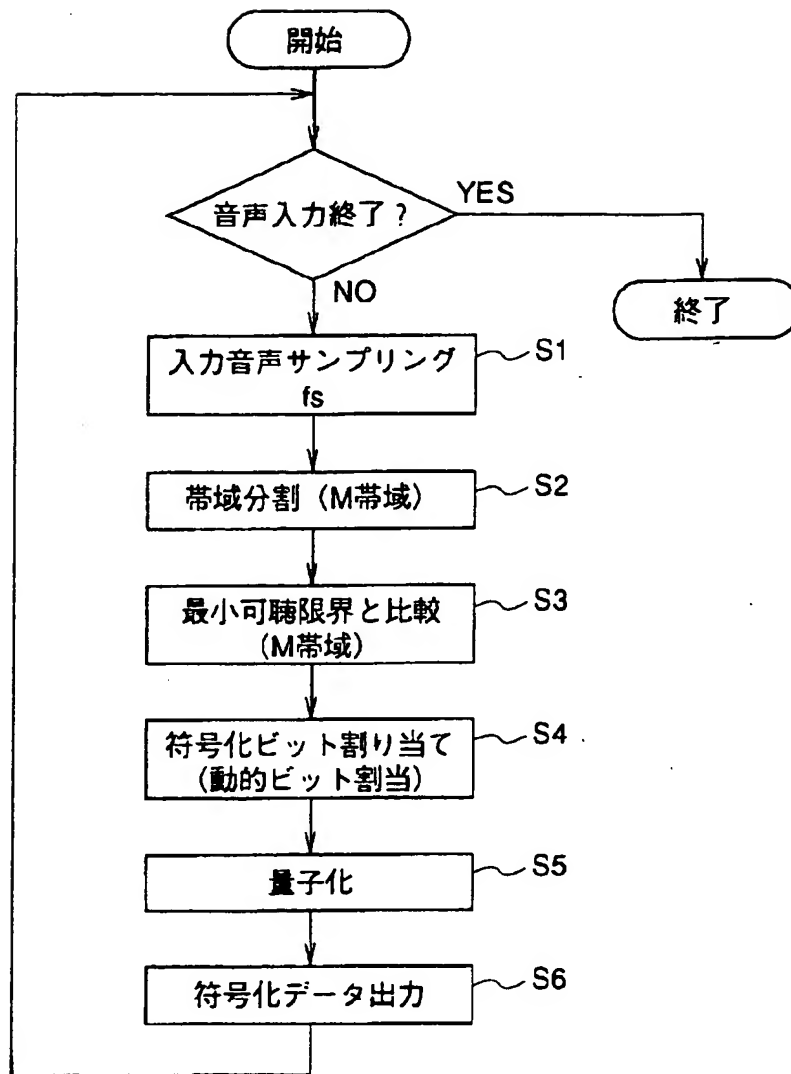
【図45】



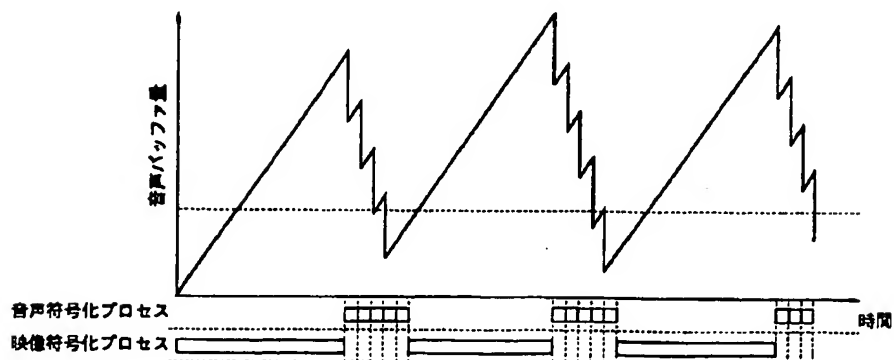
【図61】



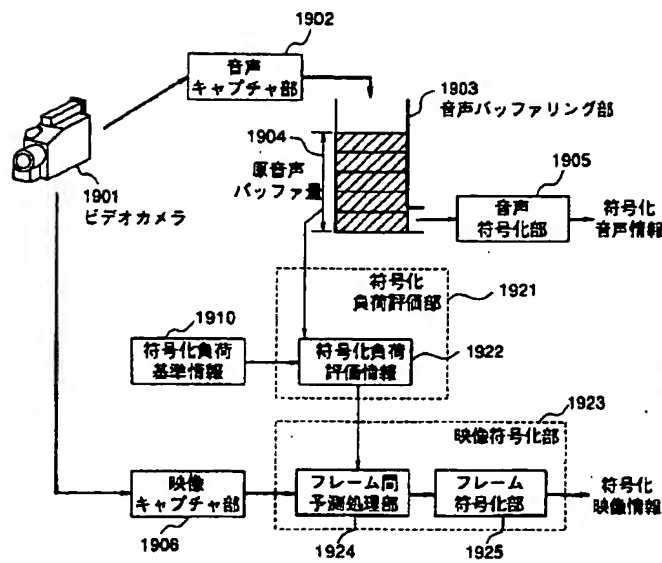
【図43】



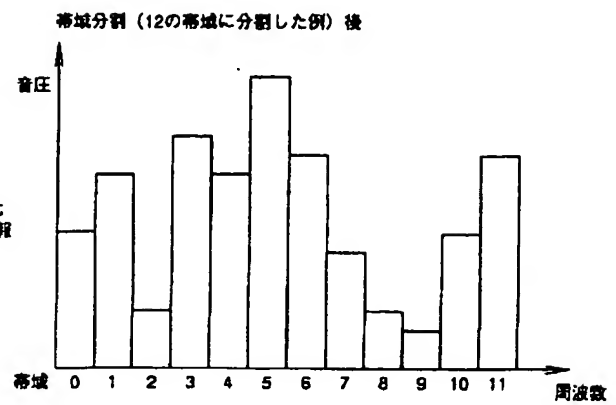
【図51】



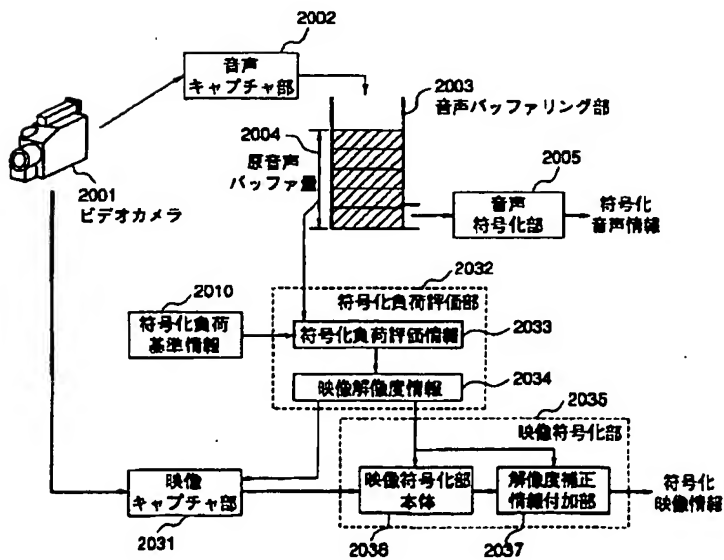
【図47】



【図63】

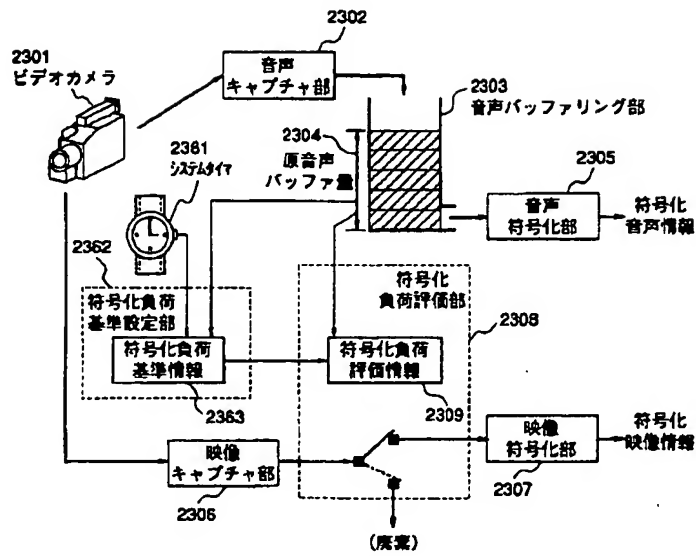


【図48】

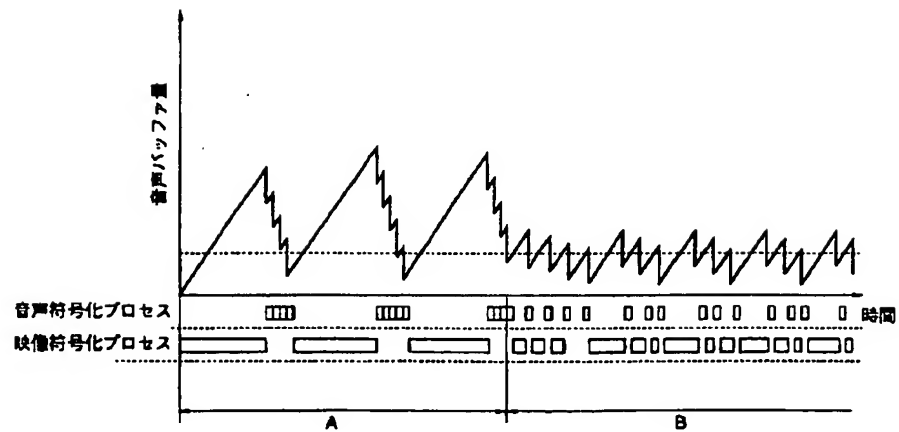


[illegible][illegible]

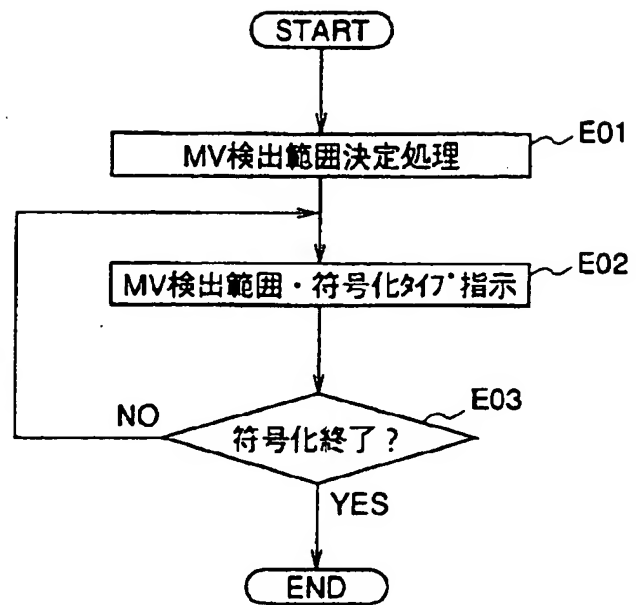
【図52】



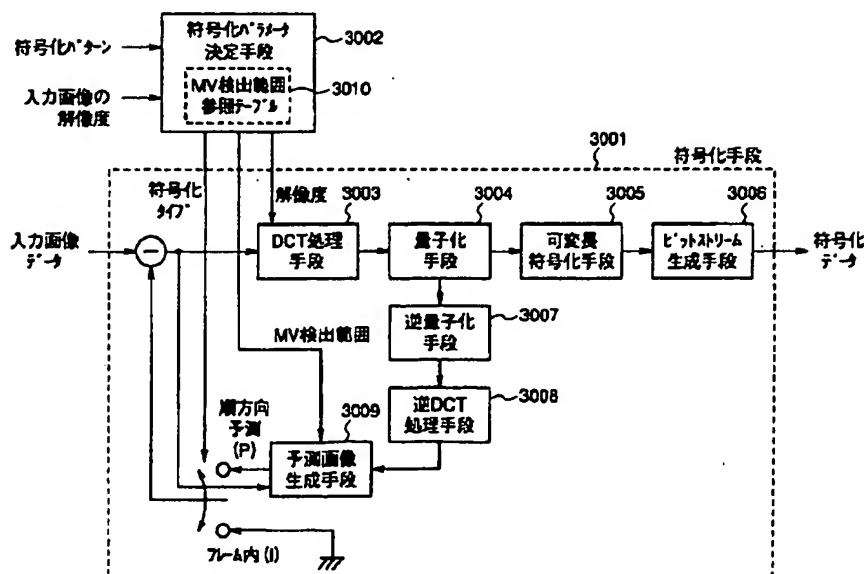
【図53】



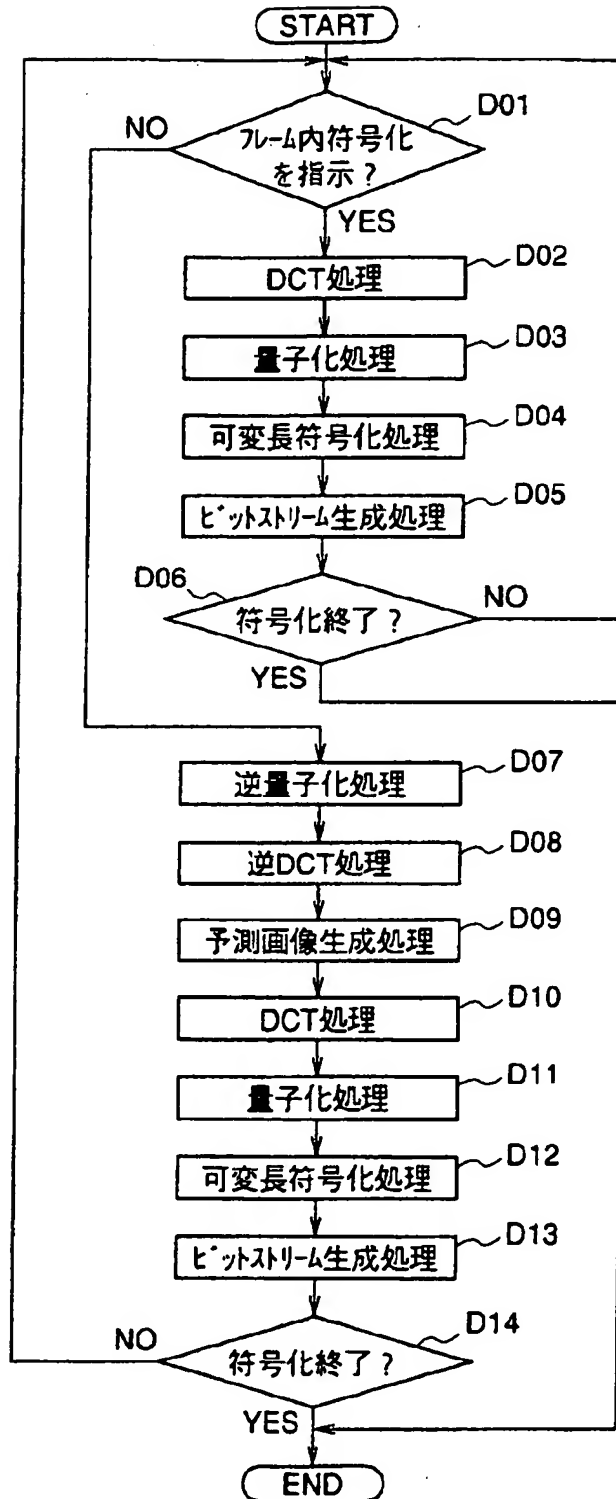
【图 5 7】



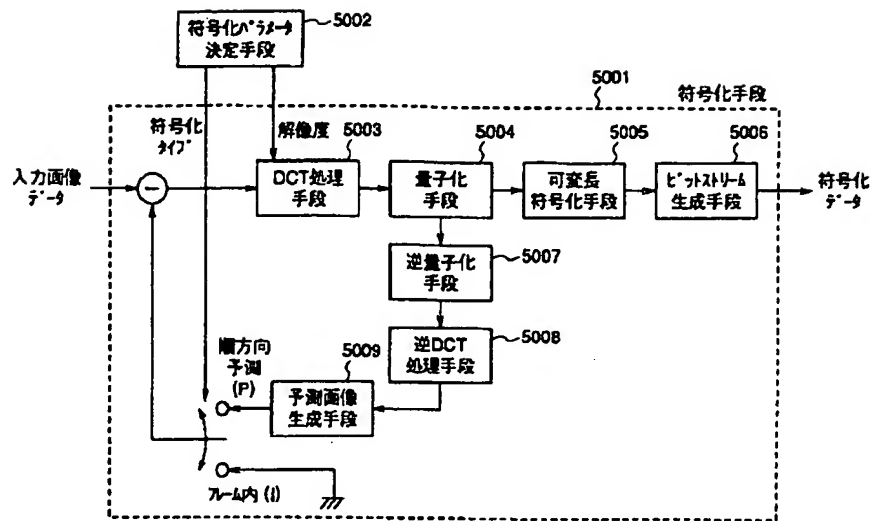
【图 5 5】



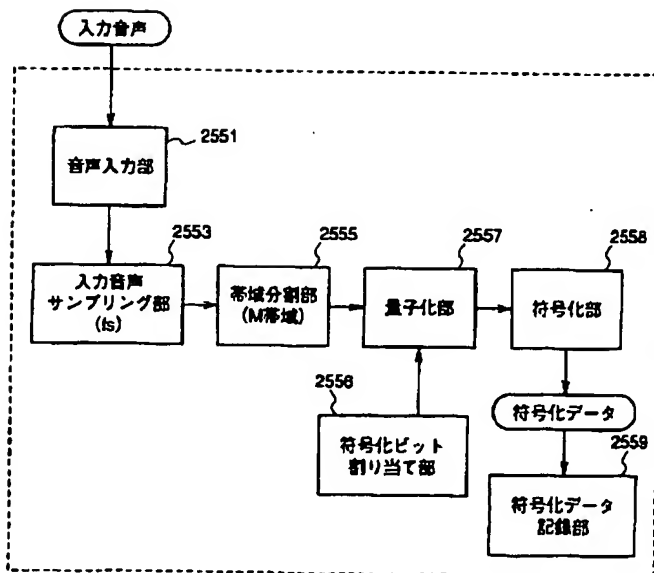
【図56】



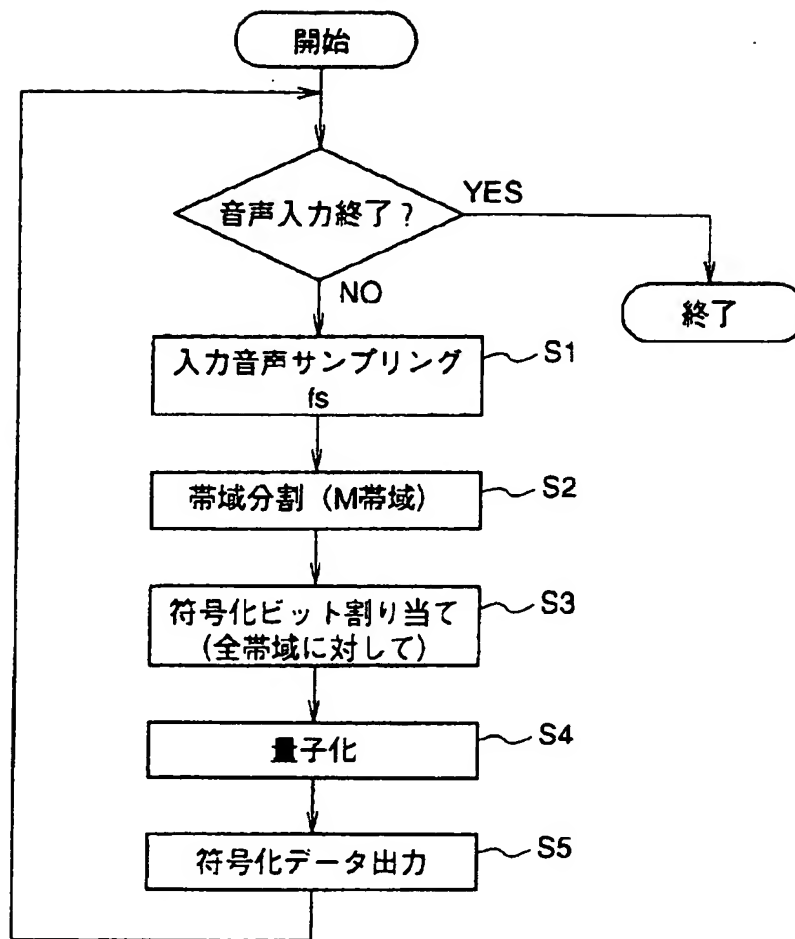
【図58】



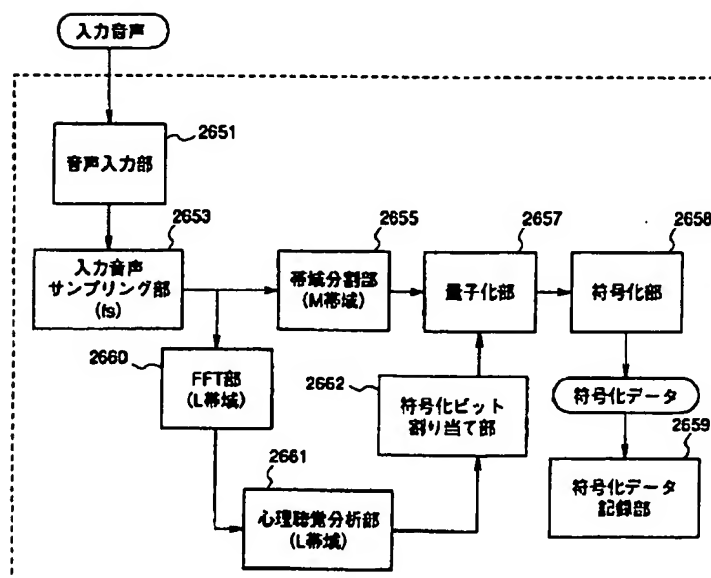
【図59】



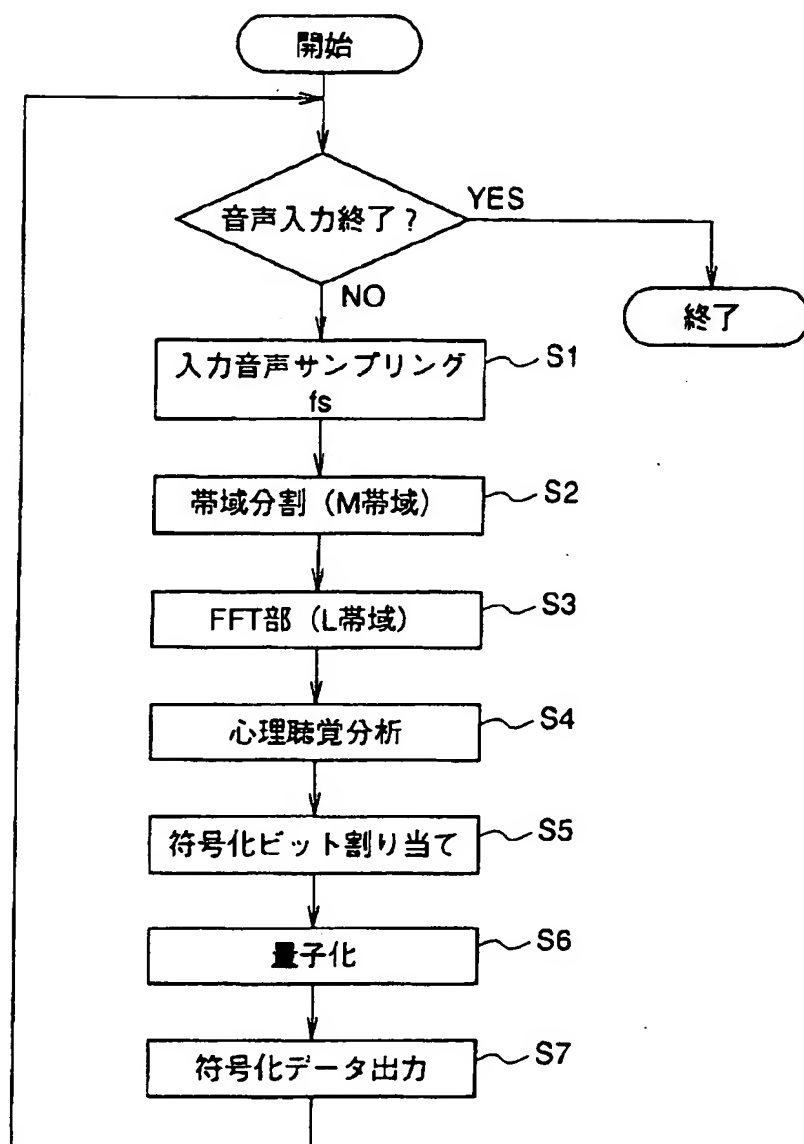
【図60】



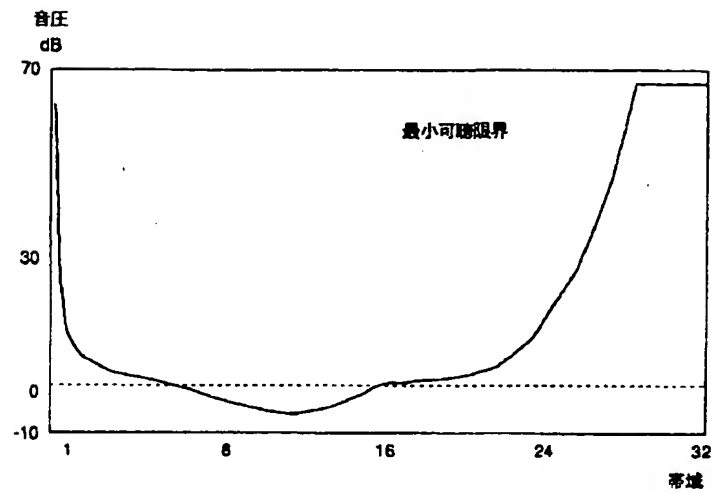
【図64】



【図65】



【図66】



フロントページの続き

(31)優先権主張番号 特願平9-42051

(32)優先日 平9(1997)2月26日

(33)優先権主張国 日本(JP)

(72)発明者 辰巳 英典

広島県広島市東区光町1丁目12番20号 株式会社松下電器情報システム広島研究所内

(72)発明者 河原 栄治

広島県広島市東区光町1丁目12番20号 株式会社松下電器情報システム広島研究所内

(72)発明者 荒瀬 吉隆

広島県広島市東区光町1丁目12番20号 株式会社松下電器情報システム広島研究所内